

Chapter 7

Mass Spectrometry

Igor A. Kaltashov and Cedric E. Bobst

Abstract Mass spectrometry is now an indispensable tool in the armamentarium of molecular biophysics, where it is used for tasks ranging from covalent structure determination to studies of higher order structure, conformational dynamics, and interactions of proteins and other biopolymers. This chapter considers the basics of biological mass spectrometry and highlights recent advances in this field (with particular emphasis on hydrogen exchange, chemical cross-linking, and native electrospray ionization mass spectrometry), evaluates current challenges, and reviews possible future developments.

7.1 Physical Principles and Instrumentation

Mass spectrometry (MS) is one of the oldest methods of instrumental analysis in chemistry, this year being the centennial of the construction of the first mass spectrometric device [1]. In addition to rather mundane applications related to molecular mass measurements (as implied by its name), MS can be used for a variety of other tasks, many of which are uniquely suited to address challenging questions in molecular biophysics and structural biology. However, it was not until the advent of the two ionization techniques capable of producing ions of large and polar molecules, electrospray ionization (ESI), and matrix-assisted laser desorption/ionization (MALDI), that MS became a commonly accepted tool in the armamentarium of modern molecular biophysics.

I.A. Kaltashov (✉) • C.E. Bobst
Department of Chemistry, University of Massachusetts-Amherst,
710 North Pleasant Street, Amherst, MA 01003, USA
e-mail: kaltashov@chem.umass.edu

7.1.1 Methods of Producing Biomolecular Ions

MS is unique among the analytical techniques commonly applied to study biomolecular structure and behavior in that the actual physical measurements are carried out in vacuum or in the gas phase, where either electric field alone or its combination with a magnetic field are used to determine ionic mass-to-charge ratios (m/z). Placing a large biomolecular ion in vacuum is no trivial task, and the absence of robust methods to do so were limiting the utility of MS in the biophysical arena until the early 1990s.

7.1.1.1 Electrospray Ionization

The advent of ESI MS in the mid-1980s [2] provided a means to observe spectra of intact proteins with no apparent mass limitation, an invention honored with a Nobel Prize in Chemistry to John Fenn in 2002 [3]. Although the ESI phenomenon was known and extensively studied for over a century, and the realization of its great analytical potential in the macromolecular realm had become apparent as early as 1960s [4], the practical applications of this ionization technique were limited to small biomolecules, such as nucleobases, amino acids [5], and short peptides [6, 7]. It was not until the demonstration of the ability of ESI to generate ionic signals for protein molecules in the form suitable for MS analysis [8] that this technique rapidly gained acceptance and recognition among MS practitioners and quickly became a tool of choice in a variety of studies of biomolecular structure.

ESI is a convoluted process, whose detailed discussion is beyond the scope of this chapter. Briefly, the protein (or, generally speaking, any biopolymer) solution is sprayed at atmospheric pressure in the presence of a strong electrostatic field, which generates metastable electrically charged droplets of the solvent encapsulating the protein molecules. Such droplets undergo a series of fission events, eventually producing either solvent-free or partially solvated protein ions. A very distinct feature of the ESI process is the accumulation of multiple charges on a single protein molecule, which leads to the appearance of multiple peaks in a mass spectrum even when a single protein is present in solution (Fig. 7.1a, b). In most cases multiple charging is the result of protonation of a number of different sites within the protein molecule, although other ubiquitous charge carriers (such as Na^+ , K^+ , NH_4^+) may also contribute. A set of ion peaks, each representing the same protein molecule and differing from the rest by the extent of multiple charging, is usually referred to as a *charge state distribution*. Determination of the protein mass based on the experimentally measured charge state distribution is relatively straightforward, and can be easily accomplished using a variety of deconvolution routines even if the mass spectrum contains several overlapping charge state distributions representing different biomolecules.

Most ESI MS analyses are carried out in the positive ion mode (where biopolymer molecules are represented in mass spectra with polycationic species), but one

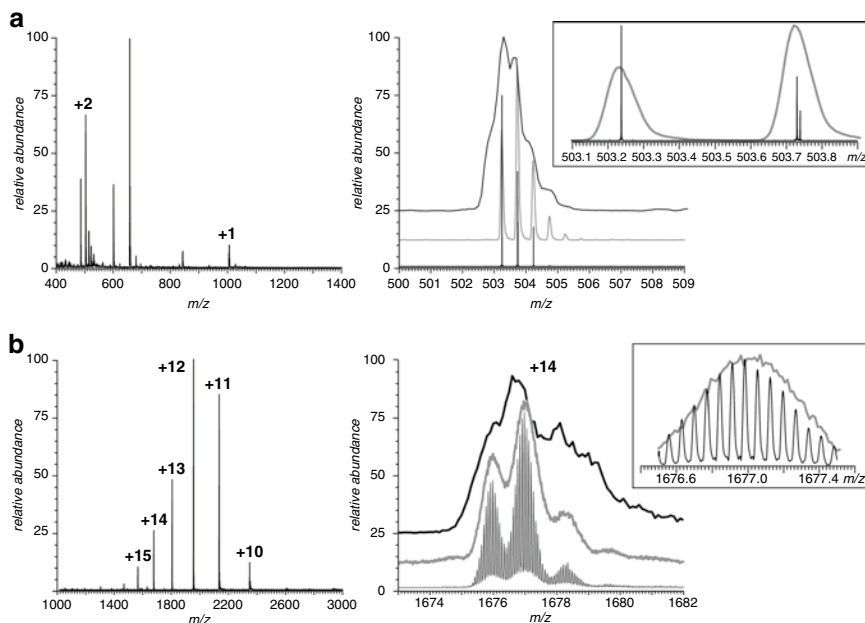


Fig. 7.1 ESI mass spectra of a peptide SWANGDEAR (**a**) and trypsin (**b**). The panels on the *left* represent full-scan mass spectra, and the panels on the *right* show detailed views of a single charge state (the three traces in each case represent mass spectra acquired with a triple quadrupole MS, hybrid quadrupole/TOF MS, and FT ICR MS), with the *insets* showing zoomed views of mass spectra acquired with quadrupole/TOF and FT ICR MS. Note that although the resolving power of TOF is sufficient to resolve isotopic peaks of the peptide ion, it fails to detect the presence of a degraded (de-amidated) form of this peptide (between m/z 503.7 and 503.8). Isotopic distribution of trypsin ions can only be resolved by FT ICR MS, although both quadrupole/TOF and FT ICR MS can resolve contributions of three different isoforms of this protein

can easily produce polyanionic species as well simply by switching the polarity of the ESI source. In this case multiple charging of macromolecules will be achieved by removing labile protons from the analyte molecule (de-protonation). While proteins are usually analyzed by ESI MS in the positive ion mode, switching to the negative ion mode could be advantageous for certain other biopolymers, such as nucleic acids. It must be stressed, however, that for any biopolymer both positive and negative ion spectra can be produced, and the charge state distributions in these spectra do not reflect the charge balance in solution [9].

7.1.1.2 MALDI

Another approach to producing macromolecular ions and transferring them to vacuum was introduced at about the same time ESI MS was developed; unlike ESI it produces ions not from the bulk of the solution, but from the interface of a

condensed phase (usually solid crystals) and the vacuum. This task is accomplished by mixing the analyte molecules with an excess of UV light-absorbing small organic molecules, which form the sample matrix, followed by irradiation with a UV laser beam. This results in rapid local heating of the matrix and subsequent ejection of a plume containing both matrix and analyte molecules from the solid surface to the gas phase and their ionization. This technique, presently known as MALDI was developed simultaneously by Koichi Tanaka [10] and Franz Hillenkamp and Michael Karas [11].

Biopolymer ions produced by MALDI can also carry multiple charges; however, the extent of protonation is significantly below that achieved with ESI. Generally, MALDI MS surpasses conventional ESI MS in terms of sensitivity and is more tolerant to salts. Superior sensitivity, relative simplicity of operation, and ease of automation have made it a top choice as an analytical technique for a variety of proteomics-related applications. On the other hand, MALDI mass spectra generally are not as reproducible as ESI mass spectra. Also, interfacing MALDI with separation techniques, such as liquid chromatography (LC), is more difficult than coupling LC to ESI MS.

7.1.2 Mass Measurements

Mass (or, more precisely, *mass-to-charge ratio*, m/z) of an ion can be determined by MS, because this characteristic of a charge-carrying particle uniquely defines its trajectory in electric (E) and magnetic (B) fields, as well as their combinations:

$$m\ddot{\vec{r}} = ze(\vec{E} + [\dot{\vec{r}} \times \vec{B}]). \quad (7.1)$$

Here ze is the ionic charge expressed as a multiple of the elementary charge e (1.6022×10^{-19} C in SI), m is its mass, while the first and second time derivatives of the trajectory vector represent its velocity and acceleration, respectively. Mass measurements are actually carried out by first separating the ions (either spatially or temporally) according to their m/z ratios, followed by detection of each type of ion, although other schemes exist where no physical separation of ions is required prior to their detection and mass measurement (*vide infra*).

The ionic m/z ratio measured by MS in most cases can be easily converted to the *ionic mass* (after taking into the account the multiple charging effect) and, ultimately, to the *molecular mass* of the analyte (after taking into the account the finite mass of the charge carriers, residual solvent, and other adducts). The notion of molecular mass (measured in unified atomic mass units, defined by IUPAC as 1/12 of the mass of a ^{12}C atom in its ground state, $u \approx 1.660\,5402(10) \times 10^{-27}$ kg) is closely related to the concept of *molecular weight*, a sum of the *atomic weights* of all atoms in a given molecule. However, the *atomic weight* of an element is a weighted average of the atomic masses of all of its stable isotopes, and the isotopic make-up is implicitly included in the definition. Contributions of isotopes are

not necessarily averaged out when ionic masses are measured by MS, and in many cases such measurements produce a distribution of masses, rather than a single value. This, of course, depends on the physical size of the analyte molecule and the mass resolution characteristics of the MS instrument (*vide infra*). Most modern MS instruments are capable of resolving isotopic distributions for relatively short peptides (Fig. 7.1a), while accomplishing the same task for proteins requires more technologically advanced (and alas, more expensive) instrumentation (Fig. 7.1b).

To avoid ambiguity in reporting molecular masses, one can use the notion of an *average mass*, which is calculated based on the entire isotopic distribution and is closely related to the *molecular weight* as used elsewhere in chemistry and related disciplines. In some applications, however, a *monoisotopic mass* would be a preferred way of reporting the molecular mass with high precision and accuracy (it is calculated based on contributions only from the lightest isotope for each element). Obviously, the use of the monoisotopic mass in reporting the MS measurement results is justified only if the resolution is high enough to afford separation of isotopic peaks in the mass spectra and the monoisotopic peak is one of the most abundant peaks in the distribution.

7.1.3 Tandem Mass Spectrometry

The most attractive features of both ESI and MALDI are their ability to generate intact macromolecular ions in the form suitable for mass measurement. However, this information is not sufficient in most instances for unequivocal identification of even small peptides, let alone large macromolecules. This task requires at least some knowledge of the covalent structure, which can be obtained by inducing dissociation of macromolecular ions and measuring the masses of the resulting fragment ions. Since most proteins and peptides are linear polymers, cleavage of a single covalent bond along the backbone generates a fragment ion (or two complementary fragment ions if the charge of the precursor ion $z=2$ or higher) classified as an *a*-, *b*-, *c*- or *x*-, *y*-, *z*-type [12, 13], depending on (1) the type of the bond cleaved and (2) whether the fragment ion contains an N- or C-terminal portion of the peptide (Fig. 7.2). Ion dissociation is usually carried out following isolation of the ion of interest from other ionic species that may be present in the mass spectrum. This approach, known as *tandem mass spectrometry* or *MS/MS*, allows the fragment ion/precursor ion correlation to be established easily [14] and is indispensable for many biophysical applications of MS (*vide infra*).

The majority of tandem MS experiments employ various means of increasing internal energy of the precursor ion to induce its dissociation. Collisional activation remains the most widely used method of elevating ion internal energy [15], which typically yields *b*- and *y*-ions, although collision-induced dissociation (CID) at high energy may also lead to formation of other fragments, particularly *a*- and *x*-type (Fig. 7.3a, b). Excitation of ions leading to their dissociation can also be

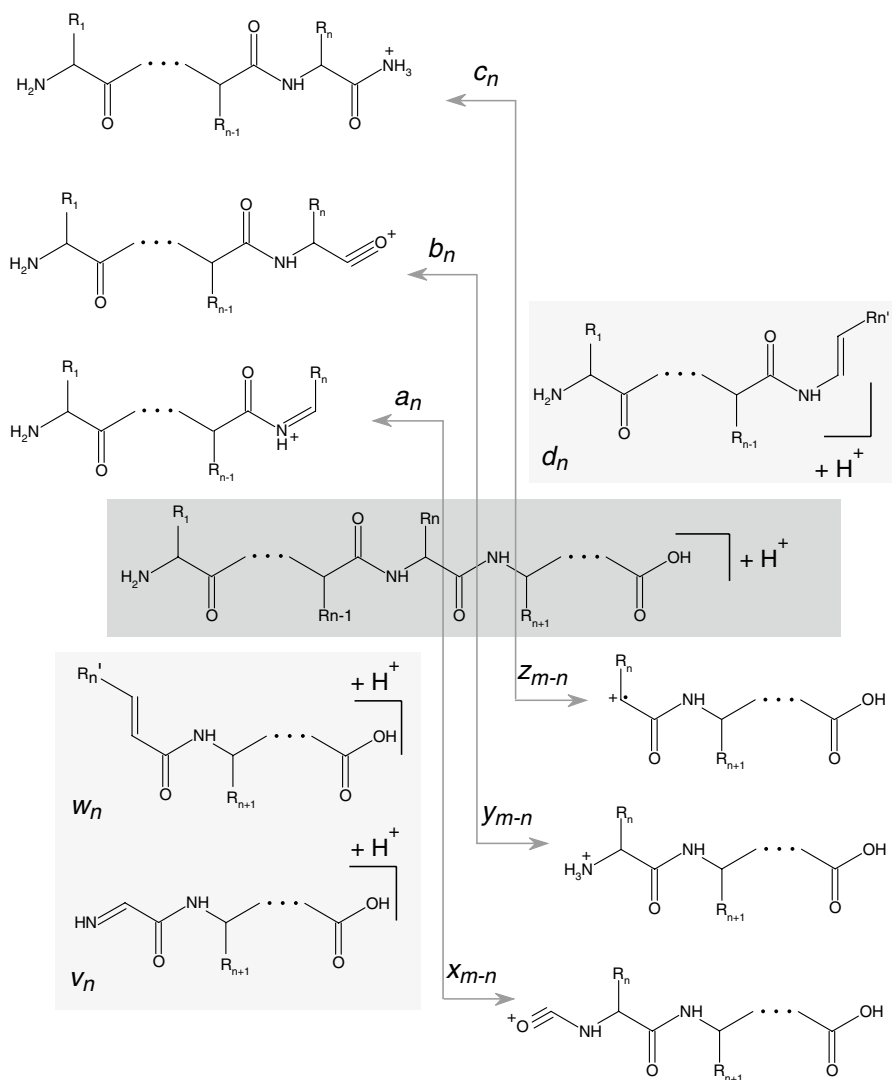


Fig. 7.2 Biemann's nomenclature of peptide ion fragments [12]. Fragment ions shown in *gray boxes* correspond to either complete or partial loss of the side chains and are usually observed only in high-energy CID

achieved using other means, such as interaction with photons (a technique known as *infrared multi-photon dissociation*, IRMPD [16]) or with electrons (two closely related techniques, known as *electron capture dissociation*, ECD [17] and *electron transfer dissociation*, ETD [18]). While the outcome of IRMPD is usually very similar to low-energy CID, ECD, and ETD typically generate *c*- and *z*-fragments, and often provide more extensive sequence coverage in polypeptides compared to

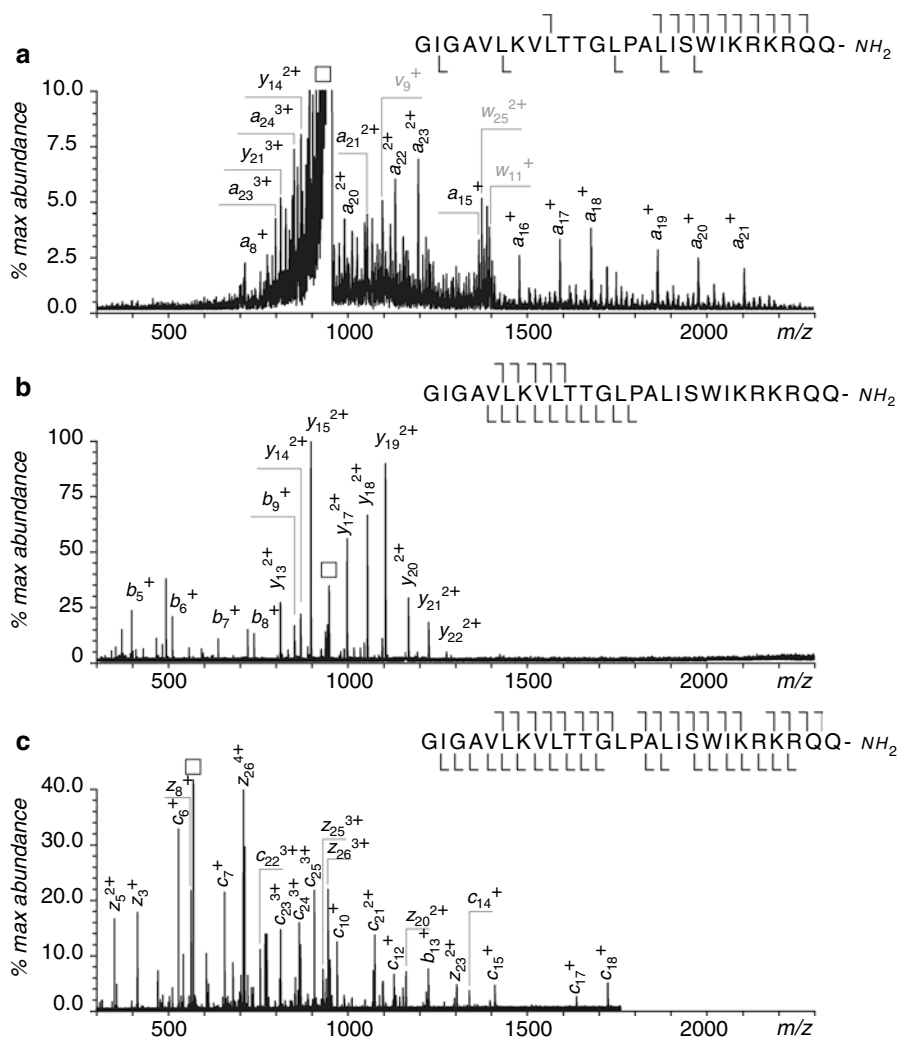


Fig. 7.3 High-energy CID (a), low-energy CID (b), and ECD (c) fragmentation spectra of a 2.8 kDa melittin peptide. Only the most abundant fragment ions are labeled in the spectra

conventional CID (Fig. 7.3c). Another very attractive feature of electron-based fragmentation techniques is their ability to preserve labile groups introduced through posttranslational modification (PTM) of proteins and cleave disulfide bonds in peptide polycations [19], a challenging task when other methods of ion activation are employed. The fragmentation patterns produced by ECD and ETD are frequently complementary to the CID-generated fragments [20], hence the benefit of using multistage fragmentation (the so-called MSⁿ experiments) consisting of both CID and ECD (or ETD).

7.1.4 Common Types of Mass Analyzers

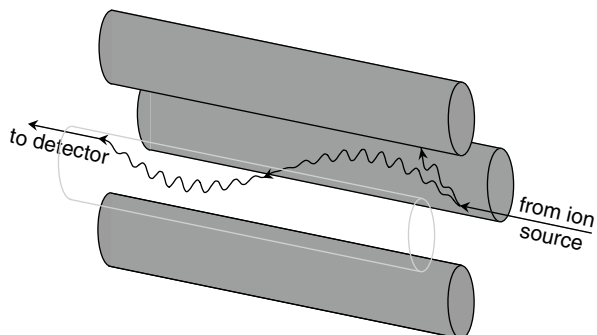
As has been already mentioned in this chapter, m/z measurements of macromolecular ions by MS rely on the unique dependence of the ionic trajectory in electric and magnetic fields on this parameter as shown in equation (7.1). The practical implementation of this principle takes a wide variety of approaches, hence a great number of mass analyzers which differ from each other not only by the amount and quality of information that can be extracted from mass measurements but also by price. Given the obvious space limitations of this volume, we cannot provide extensive coverage of all available types of mass analyzers, but instead focus our attention on three different types representing the ends and the middle of both performance and price scales. These are quadrupole, time-of-flight (TOF), and Fourier transform ion cyclotron resonance (FT ICR) mass spectrometers.

7.1.4.1 Quadrupole, Triple Quadrupole, and Ion Trap MS

Strictly speaking, *quadrupole MS* should be called a mass filter, rather than a mass analyzer, since the dynamic quadrupolar electric field employed by this device allows ions within a narrow m/z range to be transmitted through this device and eventually reach a detector, while all other ions assume unstable trajectories and are lost prior to detection (Fig. 7.4). The m/z range of a typical quadrupole MS is limited to 4,000 (with many commercial instruments having even less generous m/z limits). The mass resolution of a quadrupole MS is not constant across the m/z scale, and rarely exceeds the level of several thousands. On the other hand, these devices provide good sensitivity and are capable of obtaining mass spectra fast enough to allow direct coupling to LC. MS/MS experiments can be carried out if three quadrupoles are arranged in tandem (a configuration referred to as QqQ, or so-called *triple quadrupole MS*). The first quadrupole is set to transmit ions of certain m/z value (precursor ions), while the second is used as a collision cell and transmits all ions (precursor and CID fragments) into the third quadrupole, which is scanned to obtain a fragment ion spectrum.

Other MS/MS experiments can be designed; for example, the third quadrupole can be set to allow the transmission of fragment ions at certain m/z values, while the first quadrupole is scanned. Mass spectra acquired in this mode contain peaks of all ions whose fragmentation gives rise to a selected fragment (the so-called *precursor ion scans*). Alternatively, scanning both first and third quadrupole filters at the same rate but with a fixed m/z offset while generating fragment ions in the second nondiscriminating quadrupole produces a spectrum of ions that undergo fragmentation via loss of a specific neutral fragment (the so-called *constant neutral loss scans*). Triple quadrupole mass spectrometers are indispensable in applications that require quantitation of both small organic and biological analytes to be carried out. However, modest resolution and m/z range of such mass spectrometers limit their use in biophysical and structural biology studies, although these devices are

Fig. 7.4 A schematic representation of a quadrupole mass filter with examples of stable and unstable ion trajectories



often interfaced with other (higher-end) mass analyzers to produce hybrid mass spectrometers.

Quadrupolar devices can also be used to construct a different type of a mass analyzer, one where instead of being analyzed in a single pass through the dynamic quadrupolar field region, ions are stored (trapped) for prolonged periods of time [21]. The simplest design of such an ion trap is a segmented quadrupole (based on a triple quadrupole design), in which the central pressurized segment confines the ions radially in a dynamic (radio frequency) quadrupolar field, while the terminal segments provide repulsive DC potentials at either end that prevents the ions from escaping the central quadrupole in the axial direction. An alternative design (which is frequently referred to as a 3D ion trap to distinguish it from the linear trap described above) can be viewed as a single quadrupole filter that has been made into a toroidal device by connecting the opposite ends of each quadrupole rod and then “collapsing” this four-ring structure towards its axis of radial symmetry. In this case only one ring (the furthest from the axis) remains a ring, while the one closest to the axis completely disappears, and two other rings become endcaps flanking the remaining ring. This three-electrode system can be used to create a 3D quadrupolar electrical field, which confines ions within this device, a process that is greatly facilitated by the presence of He gas, which remove excess energy from ions via the so-called collisional damping [22, 23]. Gradual variation of electrode potentials destabilizes the trapped ions in an m/z -sensitive fashion and forces them to leave the confines of the trap, a feature that enables both MS measurements and precursor ion selection for MS/MS experiments; this field-induced external excitation can also be used to ramp-up the energy of the ions, which is then converted to internal energy upon collisions with He atoms, and eventually leads to ion dissociation [22–25].

A very significant advantage of both types of ion trapping devices described above over their progenitor quadrupole MS is that MS/MS measurements can be carried out within a single analyzer, without the need to have a dedicated collision cell and a second mass analyzer. Furthermore, any of the fragment ions produced in the course of an MS/MS experiment can also be isolated in the trap, collisionally

activated and fragmented, followed by the acquisition of a mass spectrum of the second generation of fragment ions. This process can be repeated any number of times, as long as the number of ions remaining in the trap is high enough to provide a usable signal-to-noise ratio. Such experiments are referred to as *multi-stage tandem MS*, or simply *MSⁿ*. Due to significant improvements in the performance of ion traps in the past 2 decades, ease of operation and relatively low cost, they have become very popular, both as standalone mass spectrometers and as part of hybrid instruments. Limitations of ion traps are similar to those of quadrupole MS: modest mass resolution and relatively low upper limit of the *m/z* range where MS (and MS/MS) data can be collected.

7.1.4.2 TOF MS and Hybrid Quadrupole/TOF MS

Ion separation in the *TOF MS* is based on the fact that the velocity *v* of an ion accelerated in an electrostatic field will be determined by the magnitude of the acceleration potential U_0 and the ionic *m/z* ratio. Measuring the time period needed to traverse a field-free drift region of length *D* would then allow the ionic *m/z* ratio to be determined:

$$t = \frac{D}{v} = \sqrt{\frac{m}{2zeU_0}} \cdot D \quad (7.2)$$

This approach, however, results in relatively poor mass resolution, mostly due to a significant spread of ionic kinetic energies prior to acceleration. To correct this, several approaches can be used, where energy focusing of the ions is done by delaying ion acceleration using pulsed (delayed) extraction [26] or by using the so-called ion mirror or *reflectron* [27]. The principle of the reflectron operation is illustrated in Fig. 7.5: if two identical ions have different velocities, the faster ion will penetrate deeper into the decelerating region of the reflectron, and its overall trajectory path will be longer. After its reemergence from the reflectron, this ion would still have a higher velocity, but it will be lagging behind the slower ion due to spending longer time in the decelerating region. Such relatively simple single-stage reflectrons can only perform first order velocity focusing, but more sophisticated devices (e.g., double stage ion mirrors) can provide velocity focusing to a higher order [28, 29].

Reflectrons also allow MS/MS measurements to be carried out with a single TOF mass analyzer [28], although a combination of two TOF analyzers or a hybrid instrument consisting of TOF and another, lower resolution, mass analyzer (such as a quadrupole MS) usually offer more flexibility in experiment design and deliver better data quality. A hybrid quadrupole-TOF instrument is a particularly popular configuration, which is offered by several manufacturers of MS instrumentation. Typically, a front-end quadrupole is used for mass-selection of precursor ions, followed by an RF-only quadrupole serving as a collision cell, the fragment ions are then analyzed with high resolution by a reflectron-equipped TOF section of the

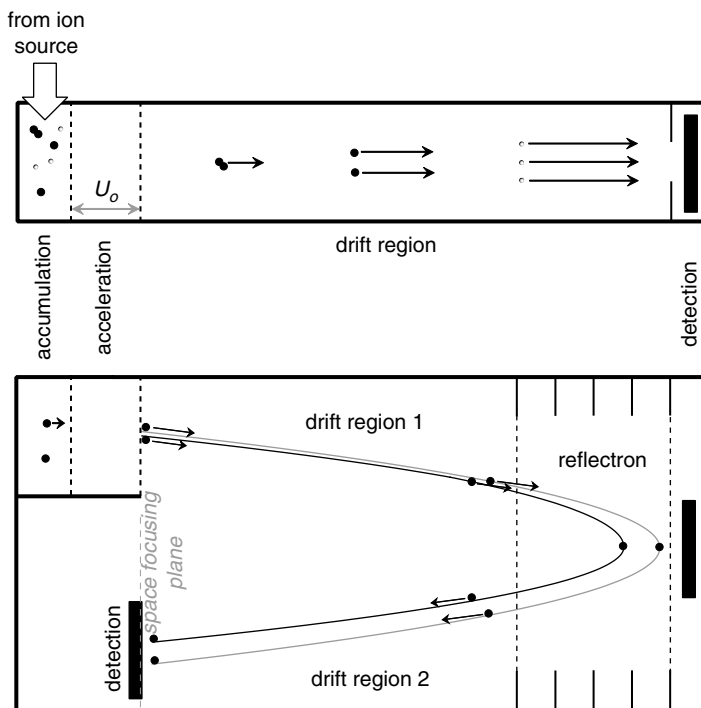


Fig. 7.5 Schematic diagrams of linear (*top*) and single-stage reflectron (*bottom*) time-of-flight mass spectrometers

instrument. MS1 measurements are carried out by operating both quadrupole segments in the RF-only mode, so that they only serve as ion guides; all mass measurements are carried out by the TOF analyzer, which offers both better resolution ($>10,000$) and m/z range vastly superior to that of the quadrupole MS.

7.1.4.3 FT ICR MS

FT ICR MS is an example of a high-performance mass spectrometer employing an ion trapping mass analyzer. However, unlike its relatively inexpensive cousins, the quadrupolar ion trap and linear ion trap considered in Sect. 1.4.1, it offers unparalleled mass resolution and unmatched mass accuracy (another high-performance mass analyzer based on the ion trapping principle is the orbitrap MS [30, 31]). Ion trapping is achieved in FT ICR MS by using a combination of electrostatic and magnetic fields, as shown in a schematic form in Fig. 7.6. A DC potential applied to the front and back plates of the cubic cell restricts the ionic motion along the z -axis, essentially locking the ions in the cell following their injection from the external source. A strong magnetic field (typically 4.7–12.0 T) applied in the direction of the

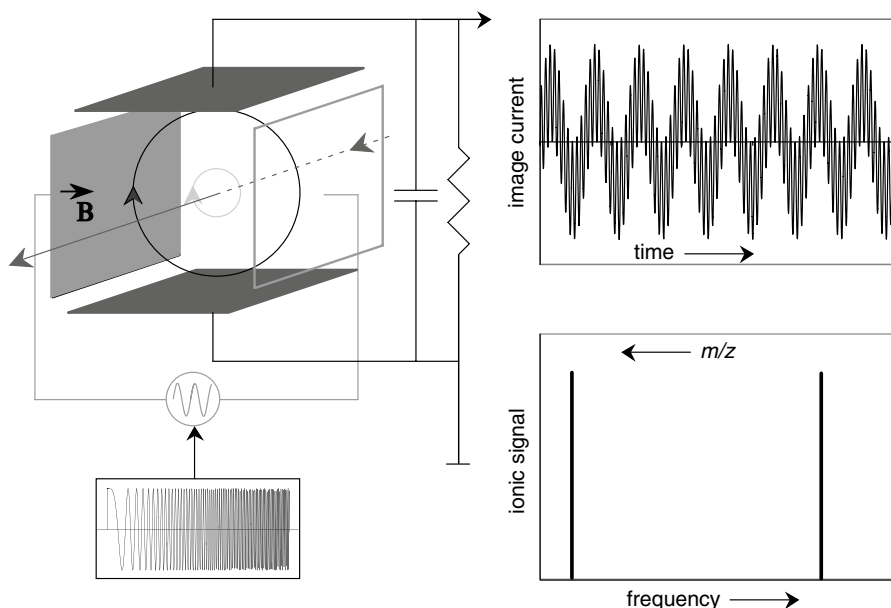


Fig. 7.6 Principal of ion trapping, broadband excitation and detection in FT ICR MS. Reproduced with permission from [132]

z -axis exerts a Lorentz force on the trapped ions, which acts as a centripetal force, inducing a circular (cyclotron) motion in the (x, y) plane. The frequency of the cyclotron motion ω_c is independent of the ionic energy, but is uniquely determined by its m/z ratio and the strength of the magnetic field B :

$$\omega_c = \frac{zeB}{m}, \quad (7.3)$$

providing the physical basis of the mass measurement. Since frequency is a physical parameter that can be measured very accurately, mass spectrometers based on the principle of cyclotron motion can provide the highest accuracy in m/z measurements.

Ion detection in FT ICR MS is done by measuring the magnitude of the image current induced on the detection plates by the ion orbiting in the space between them (Fig. 7.6). Since unsynchronized motion of a large number of ions generates zero net current, ion detection must be preceded by ion excitation (e.g., by applying a uniform harmonic electric field in the direction orthogonal to the magnetic field). If the field frequency is the same as the cyclotron frequency of the orbiting ions, they will be synchronized (brought in phase with the field). Such resonant excitation also elevates ion kinetic energy, increasing the radii of their orbits, which leads to the increase of the image current induced by each ion. Synchronized ions of the same m/z ratio induce an image current, whose angular frequency ω is equal to their

cyclotron frequency ω_c and the current amplitude is proportional to the number of ions in the cell [32]. If several types of ions (with different m/z ratios) are present in the cell, their excitation/synchronization requires application of a broadband chirp as opposed to a harmonic signal, and the resulting image current is a superposition of several sinusoidal signals (the actual cyclotron frequency in a real ICR cell is lower than ω_c due to the influence of a trapping electrostatic potential). Fourier transformation of such a spectrum allows the cyclotron frequencies of all ions to be determined and their m/z values calculated (Fig. 7.6).

Apart from ultra-high mass resolution and accuracy, a great advantage offered by FT ICR MS over most other mass analyzers is that it allows all ions across a wide m/z range to be detected (1) simultaneously within a very short period of time and (2) in a nondestructive fashion. The latter feature allows the data acquisition to be carried out with the same ion population over an extended period of time using multiple remeasurements, forming the basis of the MS^n (as opposed to MS/MS or MS^2) experiments. Ion isolation in the ICR cell can be achieved using inverse FT (from the frequency to the time domain), and fragmentation of the isolated ions can be induced by either collisional activation or electron capture (other methods of ion activation, such as IRMPD, are also available). Combining FT ICR MS with another mass analyzer (e.g., quadrupole) as a front end leads to further expansion of the repertoire of the ion fragmentation techniques, e.g., by allowing ETD to be carried out under conditions of relatively high pressure prior to introduction of fragment ions to the ICR cell for either high-resolution mass analysis or interrogation with orthogonal ion fragmentation techniques that can be performed in the high-vacuum environment of the ICR cell. Combination of several ion fragmentation techniques in one experiment often provides significant improvement of the sequence coverage of macromolecular ions [33].

7.2 Analysis of Covalent Structure

7.2.1 Covalent Structure of Polypeptides and Proteins

Tandem mass spectrometry provides the means to obtain information on covalent structure of polypeptides and proteins by employing a combination of various MS-based techniques. Typically, these are grouped in two broad categories, the so-called *bottom-up* and *top-down* approaches, which are considered in the following sections.

7.2.1.1 Polypeptide Sequencing: The Bottom-Up Approach

The classical approach to polypeptide sequencing by MS relies on enzymatic cleavage of a protein to relatively short peptides, followed by their separation by LC and

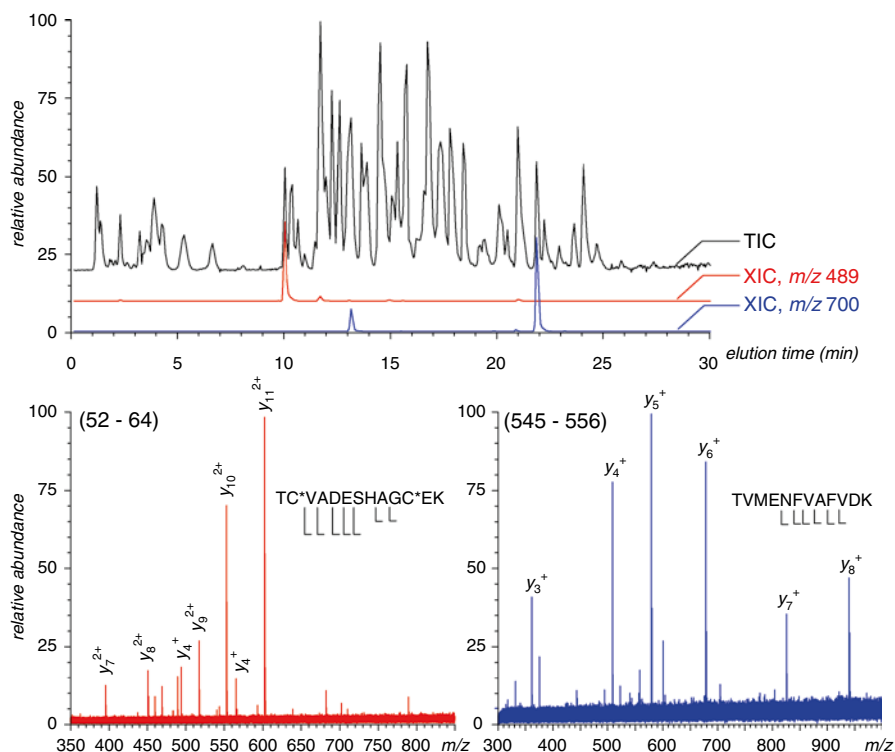


Fig. 7.7 An example of using LC/MS/MS to obtain protein sequence information. A purified 66 kDa protein bovine serum albumin has been digested with trypsin, followed by separation of proteolytic fragments on a reversed-phase (C_{18}) column with online ESI MS detection. The *black* trace in the *top* panel shows the total ionic signal recorded by ESI MS as a function of the elution time, while the *red* and *blue* traces represent ionic signals at two specific m/z values, which correspond to two proteolytic peptide ions, TCVADESHAGCEK⁶⁴ (charge state +3; both cysteine side chains are fully reduced and methylated) and TVMENFVAFVDK⁵⁵⁶ (charge state +2). The MS/MS spectra of these two peptide ions acquired in a data-dependent fashion (by selecting the most abundant ion in MS1 spectrum as a precursor for CAD) are shown in the *bottom* panels. All structurally diagnostic ions are labeled in the mass spectra, and the corresponding backbone cleavage positions are shown within each peptide's sequence

analysis of their structure using MS/MS methods [34]. The chromatographic step is usually combined with MS and/or MS/MS analysis, which frequently allows a great wealth of sequence information to be obtained in a single LC/MS/MS experiment (Fig. 7.7). The entire procedure can be automated on most commercial instruments, which allows MS/MS operation to be performed in a data-dependent fashion, while the data interpretation step is frequently assisted by database searches. The latter allows peptides and proteins to be identified even if the fragmentation patterns contain significant sequence gaps.

7.2.1.2 Polypeptide Sequencing: The Top-Down Approach

The top-down approach to polypeptide and protein sequencing completely bypasses the enzymatic degradation step, with all structural information derived from dissociation of the intact protein or polypeptide ion in the gas phase [35]. While this approach has been used successfully by many groups to obtain sequence information on relatively small proteins (<30 kDa), its application to larger proteins is not straightforward even when high-end instrumentation is used. Nevertheless, successful utilization of this methodology was demonstrated for identification of proteins beyond 500 kDa [36], although such examples remain very rare.

7.2.1.3 Posttranslational Modifications

Analysis of PTM of proteins is another area where MS-based methods of analysis are now playing a major role. Due to the labile nature of many PTMs, application of traditional MS/MS approaches to identify specific modifications and localize them within the protein sequence meets only with limited success. For example, collisional activation of glyco- and phospho-peptides frequently leads to facile removal of PTM moieties prior to cleavage of the peptide backbone, leaving no mass tags on amino acid residues that were modified and making their identification a challenging task. However, the electron-based ion dissociation techniques (such as ECD and ETD) allow this conundrum to be solved, since the fragmentation events are highly localized and do not require accumulation of vibrational energy within the peptide ion over an extended period of time (as does CAD).

7.2.1.4 Covalent Structure of Other Biopolymers

While the analysis of protein covalent structure by MS-based methods gained the most recognition and is in fact the default approach to obtaining both amino acid sequence information and mapping PTMs, structural analysis of other biopolymers also benefitted enormously from recent improvements in MS hardware and methodology. For example, both MALDI and ESI MS had been used successfully to measure masses of intact RNA molecules and other nucleic acids; however, these analyses frequently present a number of challenges, mostly due to the ability of the phosphodiester backbone of nucleic acids to form adducts with alkali and alkaline earth metal cations. This typically leads to very broad peaks in mass spectra (Fig. 7.8), although extensive buffer exchange into volatile ammonium salts to displace metal cations, desalting by metal chelation or HPLC can improve the spectral quality. Sequence information can be obtained by means of MS/MS, or simply by inducing fragmentation in the ionization source, e.g., by increasing the laser power in MALDI measurements. Dissociation of nucleic acids along the phosphodiester backbone produces structurally diagnostic ions, and these fragment ion ladders (Fig. 7.9) can be used to determine the oligonucleotide sequence. This approach to

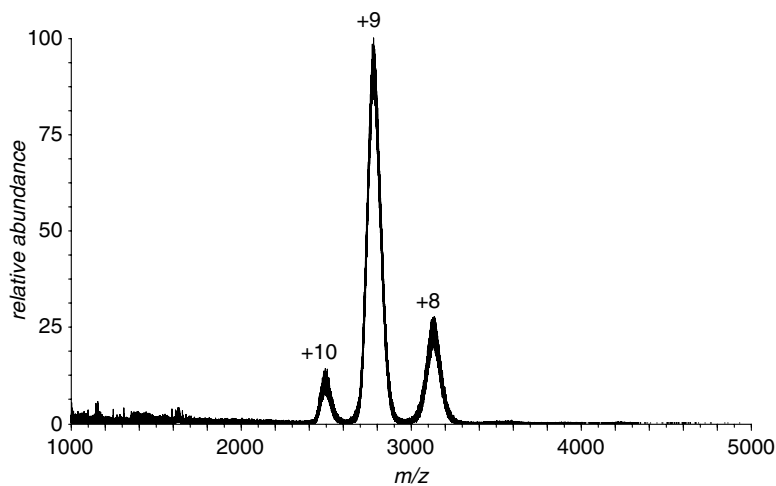


Fig. 7.8 ESI mass spectrum of $rRNA^{\text{Thr}}$ acquired with a hybrid quadrupole/TOF MS (10 μM in 20 mM ammonium acetate)

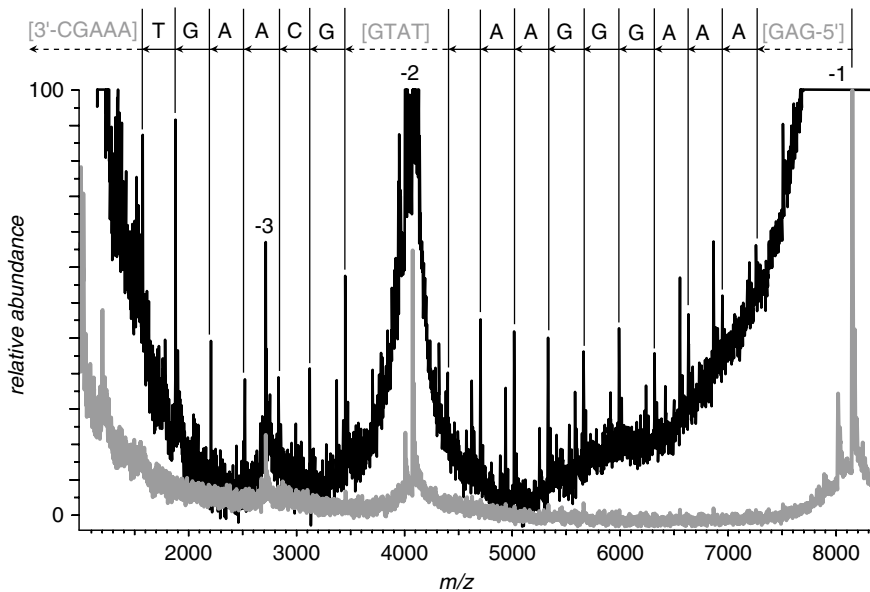


Fig. 7.9 Prompt fragmentation in MALDI MS: UV-MALDI spectra of an oligonucleotide strand acquired at increased (*top* trace) and moderate laser power. Adapted with permission from [132]

oligonucleotide sequencing is analogous to how peptide fragmentation patterns reveal the amino acid sequence (*vide supra*), although it currently remains practical only for relatively short oligonucleotides.

MS/MS methods can also be applied to obtain information on covalent structure of another type of biopolymer, polysaccharides, although these analyses tend to be less straightforward. Dissociation of polypeptide and short oligonucleotide ions tends to follow well-defined pathways, primarily occurring along the backbone. This conveniently generates structurally diagnostic fragments from which the sequence of the intact biopolymer can be derived. By contrast, dissociation of carbohydrate ions frequently leads to much more complex fragmentation patterns. Chemical bond fission commonly occurs not only between saccharide units but also across the glycosidic ring [37], and multiple rearrangement pathways are available to activated species that render analysis of tandem MS data extremely complex. Further complication arises due to the fact that unlike polypeptides and oligonucleotides, polysaccharides in general are not linear polymers, and the presence of multiple branching points makes the interpretation of MS/MS data a challenging task. Data analysis can be simplified by inducing fragmentation of polysaccharide ions with low-energy collisional activation, which typically leads to dissociation of glycosidic bonds, while leaving the rings intact. Fragmentation processes are also strongly influenced by the nature of the parent ion (alkali metal cationized species produce different fragmentation patterns compared to protonated species). Additional information can be also gained by using various chemical derivatization techniques.

Glycopeptides are another area of great interest and their structural analysis entails localization of glycosylation sites within the polypeptide chain in addition to structural studies of the carbohydrate moieties. Glycosylation site analysis is typically carried out by identifying glycopeptides among proteolytic fragments (e.g., by comparing peptide maps for intact and de-glycosylated protein). If peptide mapping of de-glycosylated protein is not feasible (e.g., due to poor solubility of the carbohydrate-free form of the protein), glycopeptides can be identified in the digest of intact glycoprotein by observing characteristic losses (e.g., 162 Da for hexose residues) in survey MS/MS spectra obtained with low-energy CID of peptide ions, since the labile nature of glycosidic bonds in the gas phase leads to their facile dissociation (*vide supra*). Precise localization of glycosylation sites can be accomplished with electron-based ion fragmentation techniques, as they preferentially cleave peptide backbone, leaving the carbohydrate chains mostly intact [38]. Complete determination of structure (especially with novel glycans) frequently requires the use of orthogonal methods, such as NMR and X-ray crystallography in addition to MS [39].

7.3 Analysis of Higher Order Structure with MS Tools

The ability of various MS-based techniques to examine covalent structure of proteins, other biopolymers and their derivatives also makes them indispensable in the studies of the higher order structure and conformational dynamics of such macromolecular

systems, which rely on various chemical probes (such as chemical labeling and cross-linking studies, to be considered later in this section). Furthermore, the unique ability of ESI to generate biomolecular ions directly from solutions kept under physiologically relevant conditions provides other opportunities to examine higher order architecture, dynamics and interactions of biopolymers, as detailed below.

7.3.1 Direct ESI MS Measurements: Characterization of Non-covalent Interactions by ESI MS

Both ESI and MALDI are rightfully credited as being soft ionization techniques, since they allow intact biopolymers to be transferred from a condensed phase to the vacuum without damaging their covalent structure. Furthermore, it was recognized soon after the introduction of these techniques into the mainstream of bioanalysis that ESI is also capable of generating ions representing intact non-covalent macromolecular complexes if the transition from solution to the gas phase is carried out under mild desolvation conditions in the ESI MS interface. The two parameters that are most critical for the survival of non-covalent complexes upon this transition are the ESI interface temperature and the electrical field in the ion desolvation region, which determines the average kinetic energy of ions undergoing frequent collisions with neutral molecules in this region. Keeping these parameters at relatively low levels allows the composition and stoichiometry of macromolecular assemblies to be determined reliably and with minimal sample consumption (Fig. 7.10). Not only can such experiments provide information on the stoichiometry of multi-protein complexes [40–44], but they may also reveal the presence of smaller ligands (e.g., metal ions and small organic molecules) within these non-covalent assemblies (see the right-hand panel in Fig. 7.10).

Reducing the efficiency of ion desolvation to ensure the survival of non-covalent complexes in ESI MS is needed in order not only to avoid collisional excitation of these species in the gas phase but also to preserve a layer of residual solvent molecules and small counterions, which are often critical for the survival of large macromolecular complexes in the gas phase [45, 46]. A frequent (and unfortunate) consequence of less-than-optimal ion desolvation in ESI MS interface is a decrease of the accuracy of mass measurements, a problem that can be dealt with very effectively by supplementing mild ESI MS measurements with those carried out under harsher conditions [47]. Although the latter step leads to partial dissociation of non-covalent complexes in the gas phase (Fig. 7.11), the surviving assemblies have lower residual solvation, and a stepwise increase of the electrostatic field in the interface region eventually results in dissociation of cofactors from the subunits, thereby allowing low molecular weight species present in each subunit to be identified and the stoichiometry established.

The ability of ESI MS to preserve non-covalent interactions has been used in the past two decades in numerous studies aimed at establishing quaternary structure of protein complexes [48]. These range from relatively modest structures to large

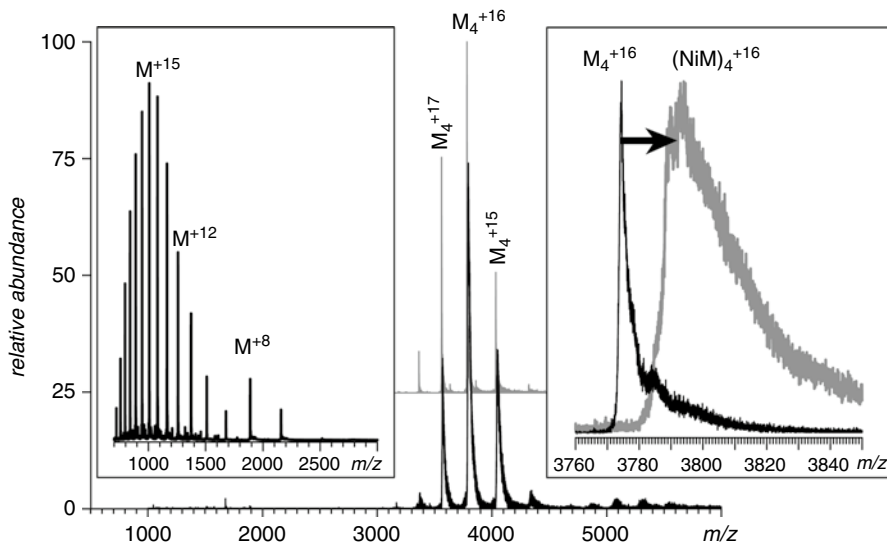


Fig. 7.10 ESI mass spectra of the apo- (gray trace) and holo- (black trace) forms of a regulatory metalloprotein NikR from *E. coli*. The main panel shows the spectra acquired under near-native conditions, when both forms assume a tetrameric structure, while the denaturing conditions (inset on the left-hand side) result in complete loss of the physiologically relevant quaternary structure and reveal only the presence of monomeric polypeptide chains. The detailed view of ionic peaks of NikR at charge state +16 (inset on the right-hand side) shows the mass difference between the ions representing the apo- and holo-forms, revealing the presence of a single metal ion in each protein subunit of holo-NikR

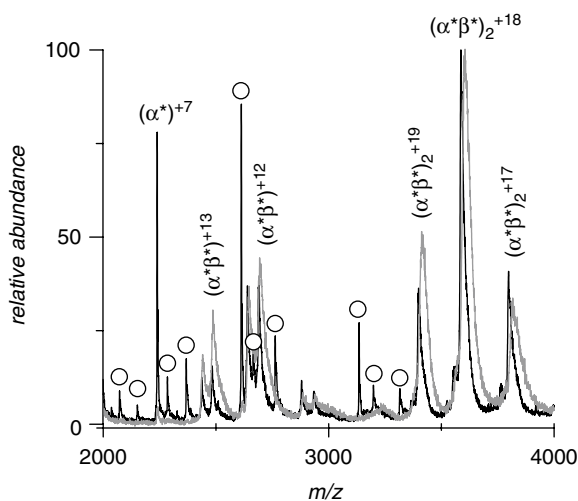
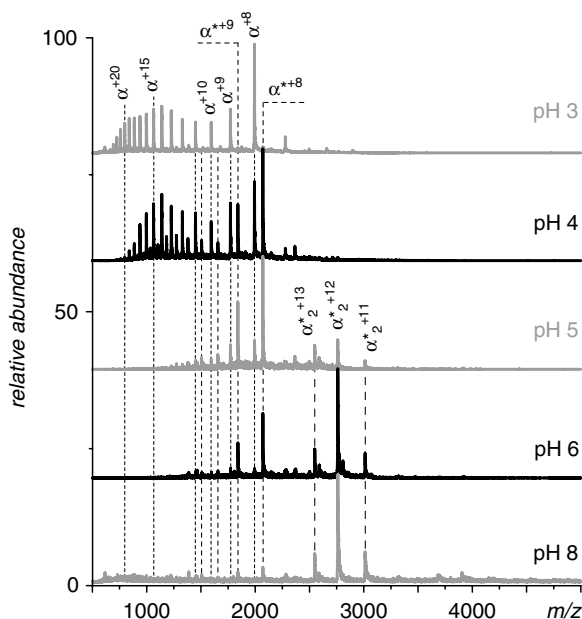


Fig. 7.11 ESI mass spectra of bovine hemoglobin acquired under very mild desolvation conditions (gray trace) to preserve all non-covalent complexes and with elevated collisional activation in the ESI interface (black trace) to enhance ionic desolvation. Note the mass shifts of ionic peaks corresponding to tetramers $(\alpha^*\beta^*)_2$ and dimers $\alpha^*\beta^*$ due to removal of a substantial fraction of residual solvent molecules. Products of gas phase fragmentation are indicated with white circles (not observed under the mild desolvation conditions)

Fig. 7.12 ESI MS monitoring of acid-induced dissociation and unfolding of homo-dimeric hemoglobin from *Scapharca* (data courtesy of Prof. Wendell P. Griffith, University of Toledo)



macromolecular assemblies whose molecular weight exceeds 1 MDa, such as intact ribosomes [49] and viral capsids [43]. This approach has also been extremely successful in probing other types of physiologically relevant non-covalent interactions, such as protein–receptor binding [50]. ESI MS can also be used to monitor changes in the composition of non-covalent associations in response to environmental factors (such as solvent composition, protein concentration, etc.). This is illustrated in Fig. 7.12 with acid-induced dissociation of dimeric hemoglobin from a mollusk *Scapharca*, where the onset of subunit dissociation clearly manifests itself via the appearance of the ionic signal representing globin monomers. Consequent dissociation of the heme group from the polypeptide chain is manifested by a mass shift of globin monomer ions corresponding to a loss of ca. 617 Da. Early stages of protein aggregation can also be monitored by ESI MS, e.g., by observing appearance of oligomeric protein ions in ESI MS in response to heat stress [51].

7.3.2 Ionic Charge State Distribution as an Indicator of Protein Compactness in Solution

So far, our discussion has been focused solely on changes of the ionic mass in ESI MS as an indicator of the changes in the protein architecture in solution. However,

careful examination of ESI MS data presented in Fig. 7.12 reveals another interesting phenomenon in addition to the dimer-to-monomer transition triggered by the acidification of the protein solution. Unlike the charge state distribution of dimer ions (α^*)₂, which remains narrow and contains only three charge states (+11, +12, and +13) as long as the dimer ions can be detected in the mass spectra, the charge state distributions of the monomer ions (both with and without the heme group, α^* and α) evolve as the solvent conditions change and become very convoluted below pH 5. The distributions of ionic charges of both of these species at pH 4 are bimodal, a feature that is usually attributed to the coexistence of two or more protein conformations in solution [9]. Native or near-native protein structures are usually very compact, and they can accommodate only a limited number of charges upon their transfer from solution to the gas phase. At the same time, even partial unfolding of a polypeptide chain results in an increase of the solvent-accessible surface area, which allows a significantly higher number of charges to be accommodated by the protein upon its transfer to the gas phase. Native and nonnative protein states often coexist at equilibrium under mildly denaturing conditions; in such situations protein ion charge state distributions in ESI MS become bimodal (as can be seen in the two top panels in Fig. 7.12), reflecting the presence of both native and denatured states. Therefore, dramatic changes of protein charge state distributions often serve as gauges of large-scale conformational changes.

The less compact the protein becomes, the higher the extent of multiple charging of the ions representing these conformers in ESI MS: as can be seen in Fig. 7.12, continuing acidification of the protein solution results in expansion of the charge state envelope of globin monomers (e.g., the mass spectrum acquired at pH 3 contains charge states +25 and higher, which are not present in the spectrum acquired at pH 4). This behavior may be indicative of the presence of several nonnative conformers in solution; however, making a distinction between the contributions made by such (partially) unfolded species to the total ionic signal is not very straightforward. Therefore, changes in the protein ion charge state distributions are frequently regarded as qualitative indicators of re- or denaturation that do not provide much information beyond loss or gain of the native fold.

This problem can be addressed at least in some cases using a procedure that utilizes chemometric tools to extract semiquantitative data on multiple protein conformational isomers coexisting in solution under equilibrium [52, 53]. Experiments are carried out by acquiring an array of spectra over a range of both near-native and denaturing conditions to ensure adequate sampling of various protein states and significant variation of their respective populations within the range of experimental conditions. The total number of protein conformers sampled in the course of the experiment can be determined by subjecting the set of collected spectra to singular value decomposition, SVD [54]. The ionic contributions of each conformer to the total signal can then be determined by using a supervised minimization routine. Application of this method to several small model proteins has yielded a picture of protein behavior consistent with that based

on the results of earlier studies that utilized a variety of orthogonal biophysical approaches [53, 55–59].

7.3.3 Hydrogen/Deuterium Exchange MS

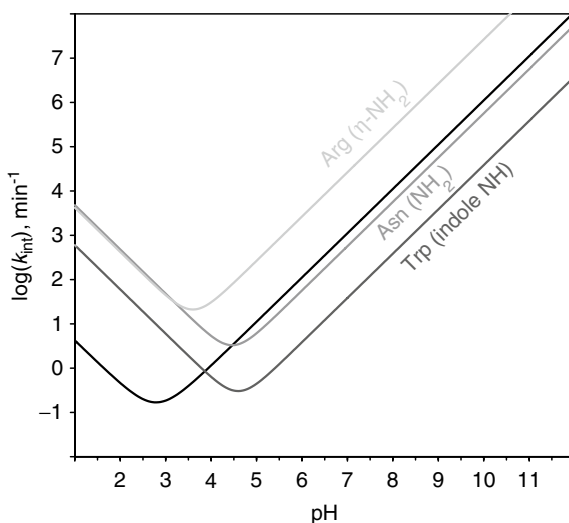
Perhaps one of the most popular and powerful MS-based experimental tools that is now widely used to study protein architecture and conformational dynamics is *hydrogen-deuterium exchange (HDX)*. The analytical value of HDX as a tool for probing macromolecular structure was recognized almost immediately after the discovery of deuterium [60] and subsequent development of the methods of production of heavy water. Initial studies of the exchange reactions between organic molecules and $^2\text{H}_2\text{O}$ carried out by Bonhoeffer and colleagues indicated that the exchange rate is very high for hetero-atoms (e.g., $-\text{OH}$ groups), while the hydrogen atoms attached to carbon atoms (e.g., $-\text{CH}_3$ groups) do not undergo exchange [61]. As early as mid-1950s, Hvidt and Linderstrøm-Lang used HDX exchange to measure solvent accessibility of labile hydrogen atoms as a probe of polypeptide structure [62, 63], and Burley et al. suggested that the extent of deuterium incorporation into a protein molecule can be measured by monitoring its mass increase [64]. However, it was not until much later that the advent of ESI and MALDI MS dramatically expanded the range of biopolymers for which the extent of deuterium incorporation could be measured by monitoring the protein mass evolution directly under a variety of conditions [65].

While MS is not the only means of detection that can be used for HDX measurements (high-resolution NMR is another popular choice), MS does offer several important advantages, namely faster time scale, tolerance to high-spin ligands and cofactors, ability to monitor the exchange in a conformer-specific fashion, as well as much more forgiving molecular weight limitations. The ability of MS to handle larger proteins and their complexes is particularly important when compared to high-field NMR, which still has limited application for proteins larger than ca. 30 kDa. Another significant advantage offered by ESI MS is its superior sensitivity, which allows many experiments to be carried out using only minute quantities of proteins.

7.3.3.1 Basic Principles of Protein HDX

HDX targets all labile hydrogen atoms (i.e., those attached to nitrogen atoms at the backbone amides and heteroatoms at polar/charged side chains), although many labile hydrogen atoms would not readily undergo HDX due to their involvement in hydrogen bonding network or sequestration from the solvent in the protein interior. Therefore, protein HDX involves two different types of reactions: (1) reversible protein unfolding that disrupts the H-bonding network and/or exposes buried segments to solvent and (2) isotope exchange at individual unprotected sites. Since protein

Fig. 7.13 Intrinsic exchange rates of several types of labile hydrogen atoms as functions of solution pH. Reproduced with permission from [132]

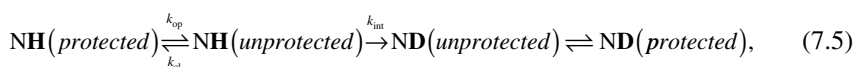


unfolding (either local or global) is a prerequisite for exchange at the sites that are protected in the native conformation, HDX reactions serve as a reliable and sensitive indicator of the unfolding events (*protection* means either involvement in the hydrogen bonding network or sequestration from solvent in the protein core). However, conformation and dynamics are not the only determinants of the HDX kinetics. Even in the absence of any protection, the exchange kinetics of a labile hydrogen atom is strongly dependent on the nature of the functional group. Furthermore, the exchange rate is strongly influenced by a variety of extrinsic factors, most notably solution pH and temperature, and the intrinsic rate constant can be expressed as [66]

$$k_{\text{int}} = k_{\text{acid}}[H^+] + k_{\text{base}}[OH^-] + k_{\text{w}} \quad (7.4)$$

The pH dependence of the cumulative intrinsic exchange rate for several types of labile hydrogen atoms, calculated based on the data compiled by Dempsey [66] is presented in Fig. 7.13.

Backbone amide hydrogen atoms constitute a particularly interesting class of labile hydrogen atoms due to their uniform distribution throughout the protein sequence, which makes them very convenient reporters of protein dynamics at the amino acid residue level (proline is the only naturally occurring amino acid lacking an amide hydrogen atom). Therefore, it is not surprising that the majority of HDX MS experiments are concerned with the exchange of the backbone amide hydrogen atoms. The mathematical formalism that is often used to describe HDX kinetics of backbone amides was introduced several decades ago and is based upon a simple two-state kinetic model [67]:



where k_{op} and k_{cl} are the rate constants for the opening (unfolding) and closing (refolding) events that expose/protect a particular amide hydrogen to/from exchange with the solvent.

In most HDX studies the exchange-incompetent state of the protein is considered to be its native state. The exchange-competent state is thought of as a nonnative structure, which can be either fully unfolded (random coil) or partially unfolded (intermediate states). Alternatively, it can represent a structural fluctuation within the native conformation, which exposes an otherwise protected amide hydrogen to solvent transiently through *local unfolding* or “structural breathing” without large-scale structure loss [68, 69]. Transitions between different nonnative states under equilibrium conditions are usually ignored in mathematical treatments of HDX, since the majority of HDX measurements are carried out under native or near-native conditions.

7.3.3.2 Global HDX MS Measurements

HDX MS measurements can provide information on *global protection* by measuring the deuterium content of the entire protein, rather than the exchange kinetics of individual amide hydrogen atoms (as done by high-resolution HDX NMR). Still, interpretation of HDX MS data often utilizes the kinetic model (7.5) by making an implicit assumption that NH(protected) and NH(unprotected) represent groups of amides, rather than individual amides that become unprotected upon transition from one state to another. Two extreme cases are usually considered: a situation when $k_{cl} \gg k_{int}$ and $k_{cl} \ll k_{int}$. The former case (referred to as the *EX2 exchange regime*) is commonly observed under native or near-native conditions, when each unfolding event is very brief, and its lifetime ($1/k_{cl}$) is much shorter than the characteristic time of exchange of an unprotected labile hydrogen atom ($1/k_{int}$). In this case the probability of exchange for even a single amide during an unfolding event will be very low, and the overall rate of exchange will be defined by both the frequency of unfolding events (k_{op}) and the probability of exchange during a single opening event:

$$k^{HDX} = k_{op} \cdot (k_{int} / k_{cl}) = k_{int} \cdot K, \quad (7.6)$$

where K is an effective equilibrium constant for the unfolding reaction, which is determined by the free energy difference between the two states of the protein. The overall exchange rate constant k^{HDX} in this case is a cumulative rate of exchange, i.e., an ensemble-averaged rate of deuterium incorporation into a molecule, and is measured as a mass shift of the isotopic cluster of a protein ion as a function of HDX time.

The opposite extreme ($k_{cl} \ll k_{int}$) is observed either when the protein is placed under denaturing conditions (which dramatically decreases the refolding rate k_{cl}), or by increasing the intrinsic exchange rate (e.g., by elevating the protein solution pH—see Fig. 7.13). As a result, the lifetime of the unprotected states become long enough to allow all exposed labile hydrogen atoms to be exchanged during a single unfolding event. In this case (commonly referred to as the *EX1 exchange regime*) the exchange rate will be determined simply by the rate of protein unfolding:

$$k^{\text{HDX}} = k_{\text{op}} \quad (7.7)$$

HDX MS measurements carried out under the EX1 conditions typically give rise to bi- or multimodal isotopic distributions, where the deuterium content of each part reflects the backbone protection levels of distinct protein conformers. This gives HDX MS the unique ability to visualize and track multiple protein states that may coexist in solution under equilibrium [70].

7.3.3.3 Local HDX MS Measurements

Replacement of each hydrogen with a deuterium (or vice versa) results in a protein mass change of about 1 Da, which makes MS a very sensitive and reliable detector of the progress of protein HDX reactions. Mass measurements of proteins undergoing HDX are usually carried out following rapid acidification of the protein solution to pH 2.5–3 and lowering the temperature to 0–4 °C, which results in significant deceleration of the chemical (intrinsic) exchange rates of backbone amide hydrogen atoms (see Fig. 7.13). These conditions, known as *HDX quenching* or *slow exchange conditions*, also result in unfolding of most proteins. Since the intrinsic exchange rates of labile side chain hydrogen atoms are not decelerated as significantly as those for backbone amides, all information on the side chain protection is generally lost during this step, leaving a single HDX reporter for each amino acid residue (again, with the exception of proline residues). Another fortunate consequence of quench-induced protein denaturation is dissociation of all non-covalently bound ligands (ranging from metal cations and small organic molecules to other biopolymers) from the protein. Therefore, measuring the protein mass under these conditions provides information only on the protein conformation and stability, rather than composition of non-covalent complexes formed by the protein and its ligands. In addition to characterizing protein conformation and stability globally, the protein can be digested with an acidic protease (e.g., pepsin) under the slow exchange conditions, and MS (usually following quick desalting and fast LC separation) can be used to measure the deuterium content of each proteolytic fragment. This produces information on protein conformation and dynamics at the local level. A typical workflow diagram of an HDX MS experiment is shown in Fig. 7.14.

Spatial resolution offered by HDX MS is usually limited only by the extent of proteolysis, which (along with other sample-handling steps) must be performed relatively quickly under the slow exchange conditions to avoid occurrence of significant *back-exchange* prior to MS measurements of the deuterium content of individual peptide fragments. In general, a large number of proteolytic fragments, particularly overlapping ones, would lead to greater spatial resolution, and hence more precise localization of the structural regions which have undergone exchange. In some cases, this may allow the backbone amide protection patterns to be determined at single-residue resolution [71], although such instances remain very rare. Supplementation of enzymatic digestion with peptide ion fragmentation in the gas phase may also enhance the spatial resolution of HDX MS measurements [72], but

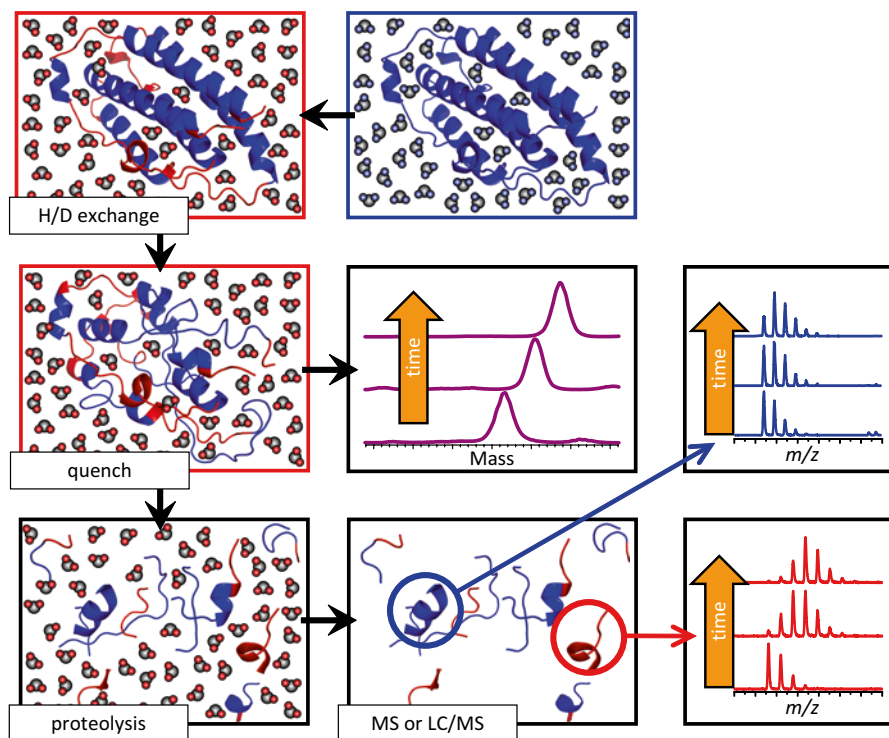


Fig. 7.14 Schematic representation of HDX MS work flow to examine protein higher order structure and conformational dynamics. The exchange is initiated by placing the unlabeled protein into a D_2O -based solvent system (e.g., by a rapid dilution). Unstructured and highly dynamic protein segments undergo fast exchange (*blue* and *red* colors represent protons and deuterons, respectively). Following the quench step (rapid solution acidification and temperature drop), the protein loses its native conformation, but the spatial distribution of backbone amide protons and deuterons across the backbone is preserved (all labile hydrogen atoms at side chains undergo fast back-exchange at this step). Rapid clean-up followed by MS measurement of the protein mass reports the total number of backbone amide hydrogen atoms exchanged under native conditions (a global measure of the protein stability under native conditions), as long as the quench conditions are maintained during the sample work-up and measurement. Alternatively, the protein can be digested under the quench conditions using acid-stable protease(s), and LC/MS analysis of masses of individual proteolytic fragments will provide information on the backbone protection of corresponding protein segments under the native conditions. Reproduced with permission from [133]

this technique has yet to be commonly accepted due to concerns over the possibility of introducing gas phase artifacts [73]. In addition to limited spatial resolution, HDX MS measurements frequently suffer from incomplete sequence coverage, especially when applied to larger and extensively glycosylated proteins. Proteins with multiple disulfide bonds constitute another class of targets for which adequate sequence coverage is difficult to achieve, although certain changes in experimental protocol can alleviate this problem, at least for smaller proteins [74]. Typically, an

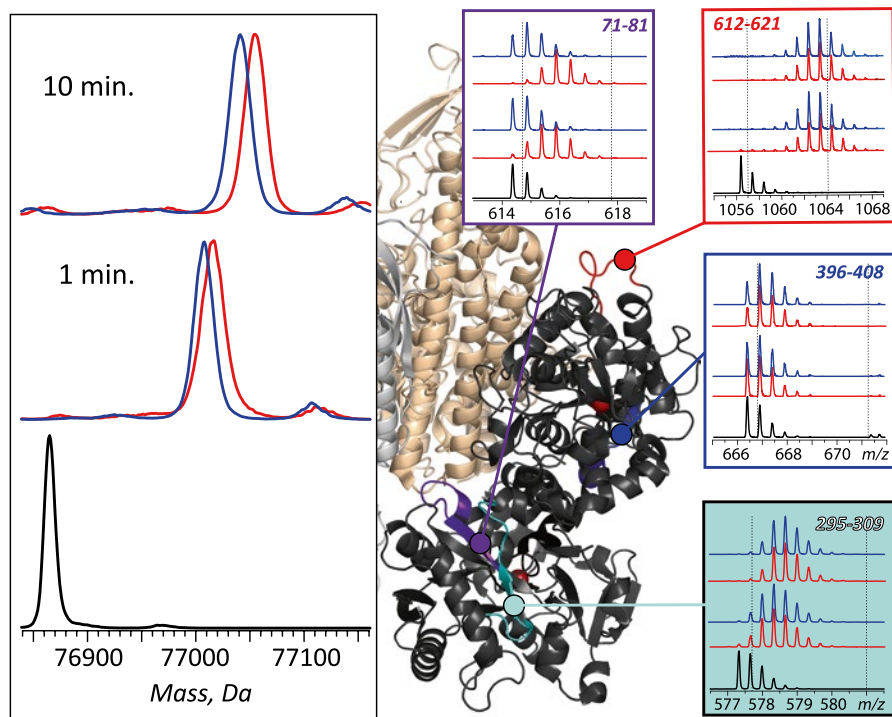


Fig. 7.15 Localization of the receptor binding interface on the surface of human serum transferrin (Tf) with HDX MS. *Left* panel: HDX MS of Tf (global exchange) in the presence (*blue*) and the absence (*red*) of the receptor. The exchange was carried out by diluting the protein stock solution 1:10 in exchange solution (100 mM NH_4HCO_3 in D_2O , pH adjusted to 7.4) and incubating for a certain period of time as indicated on each diagram followed by rapid quenching (lowering pH to 2.5 and temperature to near 0°C). The *black* trace shows unlabeled protein. *Right* panel: isotopic distributions of representative peptic fragments derived from Tf subjected to HDX in the presence (*blue*) and the absence (*red*) of the receptor and followed by rapid quenching, proteolysis, and LC/MS analysis. *Dotted lines* indicate deuterium content of unlabeled and fully exchanged peptides. Colored segments within the Tf/receptor complex show localization of these peptic fragments (based on the low-resolution structure of the complex). Adapted with permission from [73]

80 % level of sequence coverage is considered good, although significantly lower levels may also be adequate, depending on the context of the study.

An example of using HDX MS to probe protein conformation and dynamics, as well as to identify binding interface regions in a protein/receptor complex is shown in Fig. 7.15, where hydrogen exchange kinetics are measured for a diferric form of human serum transferrin (Fe_2Tf) alone and in complex with its cognate receptor. Both Tf-metal and Tf-receptor complexes dissociate under the slow exchange conditions prior to MS analysis; therefore, the protein mass evolution in each case reflects solely deuterium uptake in the course of exchange in solution (left panel in

Fig. 7.15). The extra protection afforded by the receptor binding to Tf persists over an extended period of time, and it may be tempting to assign it to shielding of labile hydrogen atoms at the protein–receptor interface. However, this view is overly simplistic, as the conformational effects of protein binding are frequently felt well beyond the interface region. The difference in the backbone protection levels of receptor-free and receptor-bound forms of Fe₂Tf appears to grow during the initial hour of exchange, reflecting significant stabilization of Fe₂Tf higher order structure by the bound receptor. Indeed, while the fast phase of HDX is often ascribed to frequent local fluctuations (transient perturbations of higher order structure) affecting relatively small protein segments, the slower phases of HDX usually reflect relatively rare, large-scale conformational transitions, such as transient partial or complete protein unfolding [75].

Evolution of the deuterium content of various peptic fragments of Fe₂Tf (right panel in Fig. 7.15) reveals a wide spectrum of protection, which is distributed very unevenly across the protein sequence. While some peptides exhibit nearly complete protection of backbone amides (e.g., segment [396–408] sequestered in the core of the protein C-lobe), exchange in many others is fast (e.g., peptide [612–621] in the solvent-exposed loop of the C-lobe). The influence of receptor binding on backbone protection is also highly localized. While most segments appear to be unaffected by the receptor binding, there are several regions where exchange kinetics are noticeably decelerated (e.g., segment [71–81] of the N-lobe, which contains several amino acid residues that form the Tf/receptor interface according to the available model of the complex based on low-resolution cryo-EM data [76]).

7.3.3.4 Local HDX MS Measurements Using a Top-Down Approach

An alternative method to probe HDX kinetics locally that does not require proteolytic fragmentation prior to MS analysis takes advantage of the ability of modern mass spectrometers to produce a wealth of structural information in tandem (MS/MS) experiments at the protein level (the top-down approach to protein sequencing discussed in Sect. 7.2.1.2). One unique advantage of the top-down HDX MS measurements that cannot be matched by the classic bottom-up type experiments is the ability to obtain protection patterns in a conformer-specific fashion. This can be accomplished by fragmenting subpopulations of protein ions, which are mass selected to include species with deuterium content representative of a certain protein conformer (this, of course, can be accomplished only under conditions favoring EX1 exchange regime in solution, so that different protein conformers can be visualized based on different levels in deuterium incorporation).

Despite the great promise of top-down HDX MS [73], applications of this technique have been limited so far due to concerns over the possibility of hydrogen scrambling accompanying dissociation of protein ions in the gas phase. Several recent studies demonstrated that the extent of scrambling is indeed negligible when ECD [77] or ETD [78] is used as a means of generating fragment ions in top-down HDX MS experiments. In addition to allowing hydrogen scrambling to be

eliminated in the top-down HDX MS experiments, both ECD and ETD appear to be superior to collisional activation in terms of generating a larger number of structurally diagnostic ions [79], allowing both better sequence coverage and enhanced spatial resolution to be achieved. In fact, in some cases it becomes possible to generate patterns of deuterium distribution across the protein backbone down to the single-residue level [77, 80].

7.3.4 Chemical Cross-Linking of Proteins

Chemical cross-linking is a classical biochemical technique used to characterize protein conformation, and it benefits tremendously from the ability of modern MS to detect and identify the products of the cross-linking reactions. Cross-linking reagents are generally classified based on their chemical specificity and the length of the *spacer arm* (cross-bridge formed between the two cross-linked sites when the reaction is complete). The chemical specificity of a cross-linker determines the overall pool of reactive groups within the polypeptide that may participate in the cross-linking reaction. Eight out of the 20 amino acid side chains are chemically reactive with good selectivity: Arg (guanidiny), Lys (ϵ -amine), Asp and Glu (β - and γ -carboxylates), Cys (sulfhydryl), His (imidazole), Met (thioether), Trp (indoyl), and Tyr (phenolic hydroxylate) [81], although virtually no reagent is absolutely group-specific.

Monofunctional (or *zero-length*) *cross-linkers* induce direct coupling of two functional groups of the protein without incorporating any extraneous material into the protein. Obviously, this becomes possible only if the two functional groups are in a very close proximity to each other, in which case the cross-linker operates as a condensing agent, resulting in the cross-linked residues becoming directly inter-joined. *Bifunctional cross-linkers*, on the other hand, contain two reagents linked through a spacer arm, thus allowing the coupling of functional groups whose separation does not exceed the spacer's length. Bifunctional reagents are further subdivided into *homobifunctional* (i.e., both cross-linking groups within the reagent targeting the same reactive groups on the protein) and *heterobifunctional* cross-linkers (coupling different functional groups on the protein).

Heterobifunctional cross-linkers may incorporate a *photosensitive* (nonspecific) *reagent* in addition to a conventional (group-specific) functionality. Such photosensitive groups react indiscriminately upon activation by irradiation. Once the specific end of such a cross-linker is anchored to an amino acid residue, the photo-reactive end can be used to probe the surroundings of this amino acid. More information on chemical cross-linkers can be found in several excellent reviews on the subject [82–85] and an outstanding book by Wong [81].

MS-assisted cross-linking studies usually aim to identify the pairs of cross-linked residues within the protein or protein complex. Such information may provide through-space distance constraints that are extremely valuable for defining both tertiary (intra-subunit cross-links) and quaternary (inter-subunit cross-links) organization of the protein when no other structural information is available. Confident

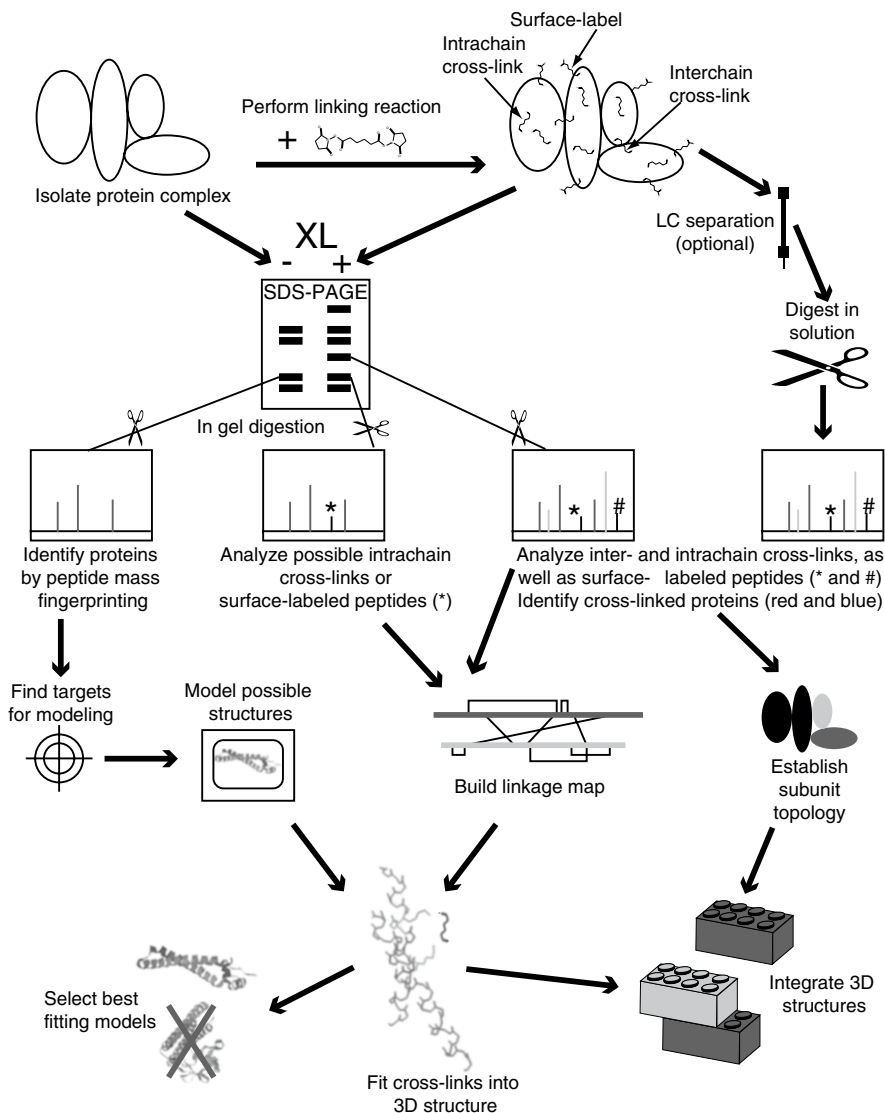


Fig. 7.16 A schematic diagram of workflow of cross-linking a multi-protein complex and integrating the levels of information into a three-dimensional model of the structure. Reprinted with permission from [86]

assignment of the pairs of coupled residues within the cross-linked protein(s) is a rather challenging experimental task. A combination of proteolysis, separation methods (e.g., LC), and mass spectrometry (and, particularly, MS/MS) provides perhaps the most elegant and efficient way of solving this problem [84, 86, 87]. Figure 7.16 shows a workflow of a typical cross-linking experiment. Separation of

proteolytic fragments prior to MS analysis usually results in significant improvements in sensitivity by eliminating possible signal suppression effects that may otherwise result in discrimination against larger (cross-linked) fragments [86]. Although peptide mapping alone can sometimes lead to confident identification of the cross-linked residues [88–90], unambiguous assignment of cross-linked peptides requires that MS/MS sequencing of the proteolytic fragments be carried out [91, 92].

As the amount of information deduced from cross-linking experiments increases, so does the complexity of data interpretation, and the tools of bioinformatics become absolutely essential to interpret the results of cross-linking experiments [93]. The task of assigning the cross-linked peptides and localizing the modification sites can be greatly assisted by a variety of automated algorithms that use MS or MS/MS data as input [86, 93]. The database mining approach to identification of cross-linked peptides mentioned earlier in this section [94] can be used even in a situation when the protein complex composition is not known a priori [95]. More sophisticated approaches, such as Xlink-Identifier [96], allow the cross-linking sites to be localized with high precision by identifying inter- and intra-peptide cross-links in addition to dead-end products and underivatized peptides. Another comprehensive cross-linking data analysis platform is MS-Bridge [97], which is part of the Protein Prospector MS data analysis suite. While these platforms were developed to support label-free analyses, several other algorithms have been developed to take advantage of isotopically tagged cross-linkers [98–101]. A comprehensive list of data analysis programs developed for interpretation of the results of cross-linking experiments can be found in a recent review article [87].

7.3.5 Chemical Labeling

Selective chemical modification [102] is another classical biophysical technique that benefitted tremendously from the recent progress in MS hardware and methodology. The unique ability of MS to localize both shielded and modified residues within a protein molecule transformed the chemical labeling technique to a highly efficient probe of higher order macromolecular structure. Most chemical modifications of an amino acid side chain alter the protein mass, hence the appeal of mass spectrometry as a readout tool for the outcome of such experiments. Interpretation of the MS and MS/MS data on chemically modified proteins is usually relatively straightforward (as compared to the analysis of cross-linked proteins) and greatly benefits from a vast arsenal of experimental tools developed to analyze PTM of proteins.

In a typical experiment, protein exposure to a certain chemical probe is followed by digestion of the modified protein with a suitable proteolytic enzyme, and mass mapping of the fragment peptides. The position(s) of the modified residue(s) within each proteolytic fragment can be reliably established using tandem mass spectrometry, as the presence of a chemical modification manifests itself as a break or a shift in the ladder of the expected fragment ions. Inter-subunit binding topology is usually determined by comparing modification patterns of the protein obtained in the

presence and in the absence of its binding partner [103], although the two experiments can be combined if the labeling agent contains a stable isotope tag [104]. An added benefit of using isotope tags is the easy recognition and quantitation of label-containing peptides and their fragments in MS and MS/MS spectra.

In addition to selective chemical labeling, protein conformation can also be characterized with *non-selective labeling*, which also offers an additional advantage of being able to determine the solvent exposure of several types of amino acids simultaneously in a single experiment. So far, the hydroxyl radical OH^\bullet is the most popular nonspecific modifier, due to its ability to induce side chain oxidation for a variety of amino acids and the relative ease of its generation in solution. Although the hydroxyl radical is relatively nondiscriminatory, and can modify virtually all types of amino acid side chains [105], the most susceptible to OH^\bullet attack are side chains containing sulfur atoms (Cys and Met), including disulfide-bonded Cys residues. The least susceptible to the OH^\bullet attack are Gly, Asn, Asp, and Ala, whose reactivity is three orders of magnitude lower than that of Cys. The great variety of OH-induced oxidation products and the large number of potential targets place a premium on the ability to detect and identify the modification sites. Usually proteolytic degradation of the modified protein followed by LC/MS and MS/MS analyses is needed in order to achieve reliable identification of oxidatively labeled amino acid side chains [105–107]. As is the case with the analysis of the results of chemical cross-linking experiments, extracting useful information from covalent labeling experimental data greatly benefits from automation [108].

One important consideration that must be kept in mind when designing or interpreting the results of both selective chemical and nonselective (oxidative) labeling experiments relates to the fact that structural information derived from such measurements is reliable only if the protein maintains its conformation during the experiment [109]. Most chemical modifications result in changing the charge of the labeled amino acid residue, and a significant alteration of the protein surface charge distribution may obviously result in conformational change. Furthermore, even the sheer size of many groups used as covalent labels may interfere with the protein's ability to maintain its conformation by creating steric constraints, but despite the extreme seriousness of this concern, less than half of all studies utilizing selective chemical labeling that were conducted in the past decade employed any means of ensuring the integrity of protein higher order structure during the experiments [109]. Artifacts associated with the influence of chemical modifications on the protein conformation can be avoided by limiting the number of modifications to one per protein molecule (in this way, reactivity of any amino acid side chain is determined only by the unperturbed protein structure [109]). While the extent of protein modification can be kept low to minimize conformational perturbations [106], this inevitably has a negative impact on the sensitivity of the measurements. A very elegant solution to this problem is based upon the realization that the extent of artifacts introduced by chemical labeling depends not only on the extent of protein oxidation but also on the time frame of the oxidation process [110]. Should this reaction time window be significantly narrow compared to the time scale of conformational changes (sub-millisecond range), the labeling pattern would reflect only the native

structure of the protein, even if the number of modified sites on each protein is significant. These considerations form the basis of a highly successful technique called fast photochemical oxidation of proteins (FPOP), where solvent-exposed amino acid residues are oxidized by OH^\cdot radicals produced by the photolysis of H_2O_2 . FPOP is designed to limit protein exposure to radicals to $<1 \mu\text{s}$ by employing a pulsed laser for initiation to produce the radicals and a radical-scavenger to limit their lifetimes [111].

7.3.6 Higher Order Structure of Other Biopolymers

7.3.6.1 DNA Higher Order Structure

Until very recently, there was substantially less interest in developing MS-based methods to probe higher order structure of DNA molecules, since they were thought to adopt only relatively few favored conformations (unlike proteins). Nevertheless, apart from the Watson–Crick double helical DNA structure (which is also known as the B-form DNA), a large number of other structures have been shown to exist, which either differ from the B conformation by arrangement of the two strands in the double helix (the so-called A and Z conformations), or by incorporating more than just two strands (e.g., triplexes and quadruplexes) [112]. Several of these non-classical DNA conformations came to prominence recently either due to their importance in designing novel therapeutic strategies [113] or for their potential use in nano-technological applications, e.g., as scaffolds of building blocks in molecular devices [114].

Similar to the studies of protein non-covalent complexes discussed in Sect. 7.3.1, ESI MS can also be used to obtain mass spectra of intact double-stranded DNA [115], as well as tetramers of short oligonucleotides that assemble to form G-quadruplex-like structures [116, 117]. Direct ESI MS measurements have also been successful as a means of monitoring DNA interaction with small ligands, most notably DNA-targeting drugs. Numerous studies have been published where this technique was employed to evaluate not only the stoichiometry of such non-covalent complexes but also their binding affinity (reviewed in [118, 119]). Information on DNA higher order structure can also be provided by using selective chemical labeling and chemical cross-linking combined with MS analysis of the products, a technique similar to those discussed in Sects. 7.3.4 and 7.3.5. While a range of chemical probes for DNA structure are available [120], mass spectrometry has not been a prominent player in this field until recently. This is beginning to change, with the realization of the enormous potential of this technique as a tool to provide rapid and sensitive characterization of the reaction products of both cross-linking [121] and chemical labeling [122].

7.3.6.2 Higher Order Structure and Dynamics of RNA

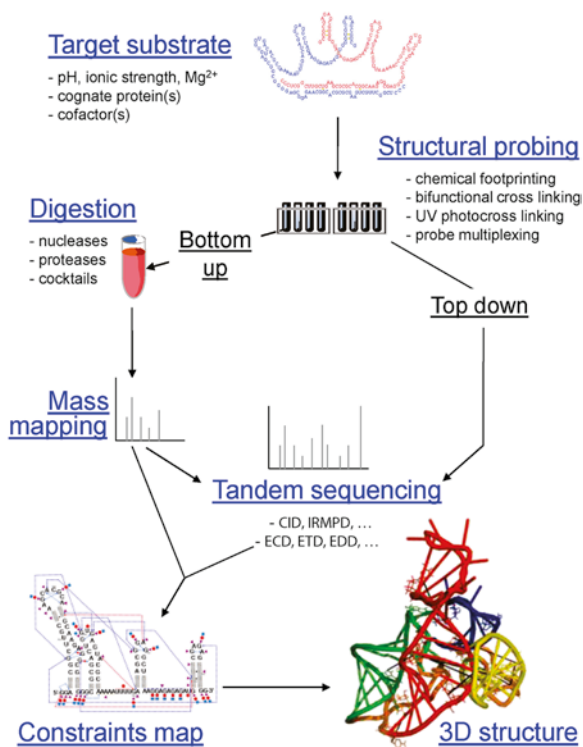
Unlike DNA, RNA molecules are known to form a rich variety of secondary and tertiary structures that make them extremely versatile, but the biophysical tools for the study of RNA structure are still somewhat less mature than those for studies of proteins. Among other things, HDX measurements have been employed to investigate structure in RNA using NMR [123, 124] and Raman spectroscopy [125]. Although the glycosidic hydrogen atoms exchange rapidly, it is possible to measure protection of the base amino and imino protons that are involved in structure, which provides information about base-pairing as opposed to bases that are involved in single stranded regions and/or bulges. While these exchange reactions are still too fast to be followed in solution by MS, hydrogen/deuterium exchange can be carried out in the gas phase, a method that shows promise for determining structural elements in oligonucleotides [126, 127].

Hydroxyl radical modification has been very successful as a means of probing oligonucleotide structure in solution, although other chemical modifications can be employed to investigate RNA structure as well. A variety of reagents are available that act as solvent accessibility probes, since they are unable to modify nucleotides involved in base-pairing, stacking, or other tertiary interactions. A similar approach can be used to probe RNA structure and RNA–protein interactions [128, 129], where the extent of chemical labeling is monitored by MS, and subsequent digestion with ribonuclease and analysis of the resulting fragments by high resolution MS allows the modification sites to be localized. In addition to solvent accessibility information, chemical labeling can also provide a measure of structural flexibility of RNA molecules [130]. Recently, a technique dubbed MS3D [92] was introduced to probe higher order structure of RNA, the workflow for which is shown in Fig. 7.17 [131]. Essentially, the structure of the polynucleotide under native conditions is probed by a series of chemical footprinting reagents. These solvent accessibility probes have varying specificity for different bases, and their reactivity is limited by the presence of base-pairing, stacking, or other tertiary interactions. Following labeling, the sites of modification are determined by a combination of bottom-up (digestion with ribonucleases) or top-down (gas phase fragmentation) methods. Additional MS/MS techniques can be used to pinpoint the labeled site to the individual nucleotide.

7.4 Current Challenges and Future Directions

Mass spectrometry has truly become a routine analytical tool in diverse fields of molecular biophysics and structural biology, although many areas remain where it still faces significant challenges. For examples, several classes of proteins are notoriously difficult to analyze using MS-based approaches, and chief among them are membrane proteins. The strongly hydrophobic or amphipathic character of membrane proteins results in their general insolubility, which makes any experimental

Fig. 7.17 General workflow for 3D-structure determination of nucleic acids based on structural probing and MS analysis (MS3D). The substrate is probed under ideal conditions preserving its native fold. Characterization of the ensuing covalent adducts can be performed under denaturing conditions, following either bottom-up or top-down approaches. The positions of probed nucleotides provide spatial constraints that are summarized on 2D maps, from which a complete, all-atom 3D structure can be readily generated through established molecular modeling protocols. Reproduced with permission from [134]



study of these proteins an extremely difficult undertaking. Mass spectrometry is not an exception, since even sequencing of membrane proteins is often problematic due to their extreme instability in solutions that are commonly used in MS work. Another obstacle to MS analysis is presented by protein aggregation, a process that is now at the cross-hair of biophysical research due to its obvious importance in the etiology of the so-called conformational diseases (such as Alzheimer's and Parkinson's), as well as its importance in the burgeoning biotechnology and biopharmaceutical sectors. Finally, mass spectrometry increasingly finds itself in the midst of the on-going paradigm shift affecting the entire field of biophysics and structural biology, namely breaking away from the reductionist description of various biophysical and biochemical phenomena, and embracing the enormous complexity of living systems. While MS in general played a very visible role in catalyzing this shift (particularly in the fields of proteomics and interactomics), many more traditional MS-based approaches to study architecture and dynamics of biological molecules were slow to respond. Clearly, biological MS is and will continue to be a very dynamic area of research, which will certainly continue to evolve and make important contributions to the Life Sciences in general, and advance the fields of biophysics and structural biology in particular.

Acknowledgements Preparation of this manuscript was supported in part by a grant from the National Institutes of Health R01 GM061666. The authors are grateful to Prof. Wendell P. Griffith (University of Toledo), Dr. Rinat R. Abzalimov (University of Massachusetts-Amherst), Dr. Guanbo Wang (Universiteit Utrecht), and Dr. Shunhai Wang (Regeneron, Inc.) for sharing unpublished data sets for this manuscript.

References

1. Thompson JJ (1912) Further experiments on positive rays. *Philos Mag* 24(140):209–253
2. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246:64–71
3. Fenn JB (2003) Electrospray wings for molecular elephants (Nobel lecture). *Angew Chem Int Ed Engl* 42(33):3871–3894
4. Dole M, Cox HL, Gieniec J (1973) Electrospray mass spectroscopy. *Adv Chem Ser* 125:73–84
5. Thomson BA, Iribarne JV, Dziedzic PJ (1982) Liquid ion evaporation/mass spectrometry/mass spectrometry for the detection of polar and labile molecules. *Anal Chem* 54:2219–2224
6. Alexandrov ML, Baram GI, Gall LN, Krasnov NV, Kusner YS, Mirgorodskaya OA, Nikolaev VI, Shkurov VA (1985) Formation of beams of quasi-molecular ions of peptides from solutions. *Bioorg Khim* 11(5):700–704
7. Alexandrov ML, Baram GI, Gall LN, Grachev MA, Knorre VD, Krasnov NV, Kusner YS, Mirgorodskaya OA, Nikolaev VI, Shkurov VA (1985) Application of a novel mass spectrometric method to sequencing of peptides. *Bioorg Khim* 11(5):705–708
8. Meng CK, Mann M, Fenn JB (1988) Of protons or proteins. *Z Phys D* 10:361–368
9. Kaltashov IA, Abzalimov RR (2008) Do ionic charges in ESI MS provide useful information on macromolecular structure? *J Am Soc Mass Spectrom* 19(9):1239–1246
10. Tanaka K, Ido Y, Akita S, Yoshida Y, Yoshida T (1987) Detection of high mass molecules by laser desorption time-of-flight mass spectrometry. In *Proceedings of the second Japan-China joint symposium on mass spectrometry*, Bando Press, Osaka, pp 185–188
11. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60(20):2299–2301
12. Biemann K (1988) Contributions of mass spectrometry to peptide and protein structure. *Biomed Environ Mass Spectrom* 16(1–12):99–111
13. Biemann K (1990) Appendix 5. Nomenclature for peptide fragment ions (positive ions). *Methods Enzymol* 193:886–887
14. McLuckey SA, Wells JM (2001) Mass analysis at the advent of the 21st century. *Chem Rev* 101(2):571–606
15. Jennings KR (2000) The changing impact of the collision-induced decomposition of ions on mass spectrometry. *Int J Mass Spectrom* 200(1–3):479–493
16. Little DP, Speir JP, Senko MW, O'Connor PB, McLafferty FW (1994) Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. *Anal Chem* 66(18):2809–2815
17. Zubarev RA, Kelleher NL, McLafferty FW (1998) Electron capture dissociation of multiply charged protein cations. *J Am Chem Soc* 120:3265–3266
18. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* 101(26):9528–9533
19. Zubarev RA (2003) Reactions of polypeptide ions with electrons in the gas phase. *Mass Spectrom Rev* 22(1):57–77

20. Savitski MM, Kjeldsen F, Nielsen ML, Zubarev RA (2006) Complementary sequence preferences of electron-capture dissociation and vibrational excitation in fragmentation of polypeptide polycations. *Angew Chem Int Ed* 45(32):5301–5303
21. Drees T, Paul W (1964) Beschleunigung von elektronen in einem plasmabetatron. *Z Chem* 180(4):340–361
22. March RE (1997) An introduction to quadrupole ion trap mass spectrometry. *J Mass Spectrom* 32(4):351–369
23. Jonscher KR, Yates JR 3rd (1997) The quadrupole ion trap mass spectrometer—a small solution to a big challenge. *Anal Biochem* 244(1):1–15
24. March RE (1998) Quadrupole ion trap mass spectrometry: theory, simulation, recent developments and applications. *Rapid Commun Mass Spectrom* 12(20):1543–1554
25. March RE (2000) Quadrupole ion trap mass spectrometry: a view at the turn of the century. *Int J Mass Spectrom* 200(1–3):285–312
26. Cotter RJ (1999) The new time-of-flight mass spectrometry. *Anal Chem* 71(13):445A–451A
27. Mamyrin BA (2001) Time-of-flight mass spectrometry (concepts, achievements, and prospects). *Int J Mass Spectrom* 206(3):251–266
28. Cotter RJ (1997) Time-of-flight mass spectrometry: instrumentation and applications in biological research. American Chemical Society, Washington, DC
29. Uphoff A, Grottemeyer J (2003) The secrets of time-of-flight mass spectrometry revealed. *Eur J Mass Spectrom* 9(3):151–164
30. Makarov A (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem* 72(6):1156–1162
31. Hardman M, Makarov AA (2003) Interfacing the orbitrap mass analyzer to an electrospray ion source. *Anal Chem* 75(7):1699–1705
32. Marshall AG, Hendrickson CL (2002) Fourier transform ion cyclotron resonance detection: principles and experimental configurations. *Int J Mass Spectrom* 215(1–3):59–75
33. Horn DM, Ge Y, McLafferty FW (2000) Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. *Anal Chem* 72(20):4778–4784
34. Fenselau C (1991) Beyond gene sequencing: analysis of protein structure with mass spectrometry. *Annu Rev Biophys Biophys Chem* 20(1):205–220
35. McLafferty FW, Breuker K, Jin M, Han X, Infusini G, Jiang H, Kong X, Begley TP (2007) Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS J* 274(24):6256–6268
36. Karabacak NM, Li L, Tiwari A, Hayward LJ, Hong PY, Easterling ML, Agar JN (2009) Sensitive and specific identification of wild type and variant proteins from 8 to 669 kDa using top-down mass spectrometry. *Mol Cell Proteomics* 8(4):846–856
37. Dorn B, Costello CE (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj J* 5:397–409
38. Mirgorodskaya E, Roepstorff P, Zubarev RA (1999) Localization of O-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Anal Chem* 71(20):4431–4436
39. Marino K, Bones J, Kattla JJ, Rudd PM (2010) A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* 6(10):713–723
40. Loo JA (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom Rev* 16(1):1–23
41. Hernandez H, Robinson CV (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* 2(3):715–726
42. Yin S, Loo JA (2009) Mass spectrometry detection and characterization of noncovalent protein complexes. *Methods Mol Biol* 492:273–282
43. Heck AJR (2008) Native mass spectrometry: a bridge between interactomics and structural biology. *Nat Methods* 5(11):927–933
44. Uetrecht C, Versluis C, Watts NR, Roos WH, Wuite GJL, Wingfield PT, Steven AC, Heck AJR (2008) High-resolution mass spectrometry of viral assemblies: molecular composition and stability of dimorphic hepatitis B virus capsids. *Proc Natl Acad Sci USA* 105(27):9216–9220

45. Abzalimov RR, Frimpong AK, Kaltashov IA (2006) Gas-phase processes and measurements of macromolecular properties in solution: on the possibility of false positive and false negative signals of protein unfolding. *Int J Mass Spectrom* 253(3):207–216
46. Freeke J, Robinson CV, Ruotolo BT (2009) Residual counter ions can stabilise a large protein complex in the gas phase. *Int J Mass Spectrom* 298(1–3):91–98
47. Lei QP, Cui X, Kurtz DM Jr, Amster IJ, Chernushevich IV, Standing KG (1998) Electrospray mass spectrometry studies of non-heme iron-containing proteins. *Anal Chem* 70(9):1838–1846
48. Heck AJ, Van Den Heuvel RH (2004) Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom Rev* 23(5):368–389
49. Gordiyenko Y, Robinson CV (2008) The emerging role of MS in structure elucidation of protein-nucleic acid complexes. *Biochem Soc Trans* 036(4):723–731
50. Leverence R, Mason AB, Kaltashov IA (2010) Noncanonical interactions between serum transferrin and transferrin receptor evaluated with electrospray ionization mass spectrometry. *Proc Natl Acad Sci U S A* 107(18):8123–8128
51. Wang G, Abzalimov RR, Kaltashov IA (2011) Direct monitoring of heat-stressed biopolymers with temperature-controlled electrospray ionization mass spectrometry. *Anal Chem* 83(8):2870–2876
52. Dobo A, Kaltashov IA (2001) Detection of multiple protein conformational ensembles in solution via deconvolution of charge state distributions in ESI MS. *Anal Chem* 73:4763–4773
53. Mohimen A, Dobo A, Hoerner JK, Kaltashov IA (2003) A chemometric approach to detection and characterization of multiple protein conformers in solution using electrospray ionization mass spectrometry. *Anal Chem* 75(16):4139–4147
54. Hendler RW, Shrager RI (1994) Deconvolutions based on singular value decomposition and the pseudo-inverse: a guide for beginners. *J Biochem Biophys Methods* 28(1):1–33
55. Borysik AJH, Radford SE, Ashcroft AE (2004) Co-populated conformational ensembles of b_2 -microglobulin uncovered quantitatively by electrospray ionization mass spectrometry. *J Biol Chem* 279(26):27069–27077
56. Frimpong AK, Abzalimov RR, Uversky VN, Kaltashov IA (2010) Characterization of intrinsically disordered proteins with electrospray ionization mass spectrometry: conformational heterogeneity of alpha-synuclein. *Proteins* 78(3):714–722
57. Testa L, Brocca S, Šamalíkova M, Santambrogio C, Alberghina L, Grandori R (2011) Electrospray ionization-mass spectrometry conformational analysis of isolated domains of an intrinsically disordered protein. *Biotechnol J* 6(1):96–100
58. Natalello A, Benetti F, Doglia SM, Legname G, Grandori R (2011) Compact conformations of α -synuclein induced by alcohols and copper. *Proteins* 79(2):611–621
59. Brocca S, Testa L, Sobott F, Šamalíkova M, Natalello A, Papaleo E, Lotti M, De Gioia L, Doglia SM, Alberghina L, Grandori R (2011) Compaction properties of an intrinsically disordered protein: Sic1 and its kinase-inhibitor domain. *Biophys J* 100(9):2243–2252
60. Urey HC, Brickedde FG, Murphy GM (1932) A hydrogen isotope of mass 2. *Phys Rev* 39:164–165
61. Bonhoeffer KF, Klar R (1934) Über den austausch von schweren wasserstoffatomen zwischen wasser und organischen verbindungen. *Naturwissenschaften* 22:45
62. Hvidt A, Linderstrom-Lang K (1954) Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochim Biophys Acta* 14(4):574–575
63. Hvidt A, Linderstrom-Lang K (1955) The kinetics of deuterium exchange of insulin with D_2O . An amendment. *Biochim Biophys Acta* 16:168–169
64. Burley RW, Nicholls CH, Speakman JB (1955) The crystalline/amorphous ratio of keratin fibres. Part II. The hydrogen-deuterium exchange reaction. *J Text Inst* 46:T427–T432
65. Katta V, Chait BT (1991) Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Commun Mass Spectrom* 5(4):214–217
66. Dempsey CE (2001) Hydrogen exchange in peptides and proteins using NMR-spectroscopy. *Prog Nucl Magn Reson Spectrosc* 39(2):135–170

67. Hvidt A, Nielsen SO (1966) Hydrogen exchange in proteins. *Adv Protein Chem* 21:287–386
68. Maity H, Lim WK, Rumbley JN, Englander SW (2003) Protein hydrogen exchange mechanism: local fluctuations. *Protein Sci* 12(1):153–160
69. Qian H, Chan SI (1999) Hydrogen exchange kinetics of proteins in denaturants: a generalized two-process model. *J Mol Biol* 286(2):607–616
70. Eyles SJ, Kaltashov IA (2004) Methods to study protein dynamics and folding by mass spectrometry. *Methods* 34(1):88–99
71. Del Mar C, Greenbaum EA, Mayne L, Englander SW, Woods VL Jr (2005) Structure and properties of *a*-synuclein and other amyloids determined at the amino acid level. *Proc Natl Acad Sci U S A* 102(43):15477–15482
72. Rand KD, Zehl M, Jensen ON, Jorgensen TJD (2009) Protein hydrogen exchange measured at single-residue resolution by electron transfer dissociation mass spectrometry. *Anal Chem* 81(14):5577–5584
73. Kaltashov IA, Bobst CE, Abzalimov RR (2009) H/D exchange and mass spectrometry in the studies of protein conformation and dynamics: is there a need for a top-down approach? *Anal Chem* 81(19):7892–7899
74. Burke JE, Karbarz MJ, Deems RA, Li S, Woods VL, Dennis EA (2008) Interaction of group IA phospholipase A2 with metal ions and phospholipid vesicles probed with deuterium exchange mass spectrometry. *Biochemistry* 47(24):6451–6459
75. Xiao H, Hoerner JK, Eyles SJ, Dobo A, Voigtman E, Mel'cuk AI, Kaltashov IA (2005) Mapping protein energy landscapes with amide hydrogen exchange and mass spectrometry: I. A generalized model for a two-state protein and comparison with experiment. *Protein Sci* 14(2):543–557
76. Cheng Y, Zak O, Aisen P, Harrison SC, Walz T (2004) Structure of the human transferrin receptor-transferrin complex. *Cell* 116(4):565–576
77. Pan J, Han J, Borchers CH, Konermann L (2008) Electron capture dissociation of electrosprayed protein ions for spatially resolved hydrogen exchange measurements. *J Am Chem Soc* 130(35):11574–11575
78. Abzalimov RR, Kaplan DA, Easterling ML, Kaltashov IA (2009) Protein conformations can be probed in top-down HDX MS experiments utilizing electron transfer dissociation of protein ions without hydrogen scrambling. *J Am Soc Mass Spectrom* 20(8):1514–1517
79. Zubarev RA, Zubarev AR, Savitski MM (2008) Electron capture/transfer versus collisionally activated/induced dissociations: solo or duet? *J Am Soc Mass Spectrom* 19(6):753–761
80. Pan J, Han J, Borchers CH, Konermann L (2009) Hydrogen/deuterium exchange mass spectrometry with top-down electron capture dissociation for characterizing structural transitions of a 17 kDa protein. *J Am Chem Soc* 131:12801–12808
81. Wong SS (1991) *Chemistry of protein conjugation and cross-linking*. CRC, Boca Raton, FL
82. Han K-K, Richard C, Delacourte A (1984) Chemical cross-links of proteins by using bifunctional reagents. *Int J Biochem* 16(2):129–145
83. Mattson G, Conklin E, Desai S, Nielander G, Savage MD, Morgensen S (1993) A practical approach to crosslinking. *Mol Biol Rep* 17(3):167–183
84. Andrea S (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom Rev* 25(4):663–682
85. Petrotchenko EV, Borchers CH (2010) Crosslinking combined with mass spectrometry for structural proteomics. *Mass Spectrom Rev* 29(6):862–876
86. Back JW, de Jong L, Muijsers AO, de Koster CG (2003) Chemical cross-linking and mass spectrometry for protein structural modeling. *J Mol Biol* 331(2):303–313
87. Jin Lee Y (2008) Mass spectrometric analysis of cross-linking sites for the structure of proteins and protein complexes. *Mol Biosyst* 4(8):816–823
88. Yang T, Horejsh DR, Mahan KJ, Zaluzec EJ, Watson TJ, Gage DA (1996) Mapping cross-linking sites in modified proteins with mass spectrometry: an application to cross-linked hemoglobins. *Anal Biochem* 242(1):55–63

89. Bennett KL, Kussmann M, Bjork P, Godzwon M, Mikkelsen M, Sorensen P, Roepstorff P (2000) Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping—a novel approach to assess intermolecular protein contacts. *Protein Sci* 9(8):1503–1518
90. Sinz A, Wang K (2001) Mapping protein interfaces with a fluorogenic cross-linker and mass spectrometry: application to nebulin-calmodulin complexes. *Biochemistry* 40(26):7903–7913
91. Yu Z, Friso G, Miranda JJ, Patel MJ, Lo-Tseng T, Moore EG, Burlingame AL (1997) Structural characterization of human hemoglobin crosslinked by bis(3,5-dibromosalicyl) fumarate using mass spectrometric techniques. *Protein Sci* 6(12):2568–2577
92. Young MM, Tang N, Hempel JC, Oshiro CM, Taylor EW, Kuntz ID, Gibson BW, Dollinger G (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc Natl Acad Sci U S A* 97(11):5802–5806
93. Leitner A, Walzthoeni T, Kahraman A, Herzog F, Rinner O, Beck M, Aebersold R (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol Cell Proteomics* 9(8):1634–1649
94. Rappsilber J, Siniosoglou S, Hurt EC, Mann M (2000) A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry. *Anal Chem* 72(2):267–275
95. Winters MS, Day RA (2003) Detecting protein-protein interactions in the intact cell of *Bacillus subtilis* (ATCC 6633). *J Bacteriol* 185(14):4268–4275
96. Du X, Chowdhury SM, Manes NP, Wu S, Mayer MU, Adkins JN, Anderson GA, Smith RD (2011) Xlink-Identifier: an automated data analysis platform for confident identifications of chemically cross-linked peptides using tandem mass spectrometry. *J Proteome Res* 10(3):923–931
97. Chu F, Baker PR, Burlingame AL, Chalkley RJ (2010) Finding chimeras: a bioinformatics strategy for identification of cross-linked peptides. *Mol Cell Proteomics* 9(1):25–31
98. Rinner O, Seebacher J, Walzthoeni T, Mueller L, Beck M, Schmidt A, Mueller M, Aebersold R (2008) Identification of cross-linked peptides from large sequence databases. *Nat Methods* 5(4):315–318
99. Gao QX, Xue S, Doneanu CE, Shaffer SA, Goodlett DR, Nelson SD (2006) Pro-CrossLink. Software tool for protein cross-linking and mass spectrometry. *Anal Chem* 78(7):2145–2149
100. Maiolica A, Cittaro D, Borsotti D, Sennels L, Ciferri C, Tarricone C, Musacchio A, Rappsilber J (2007) Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching. *Mol Cell Proteomics* 6(12):2200–2211
101. Petrochenko EV, Borchers CH (2010) ICC-CLASS: isotopically-coded cleavable crosslinking analysis software suite. *BMC Bioinformatics* 11:64–73
102. Glazer AN (1970) Specific chemical modification of proteins. *Annu Rev Biochem* 39(1):101–130
103. Hager-Braun C, Tomer KB (2002) Characterization of the tertiary structure of soluble CD4 bound to glycosylated full-length HIVgp120 by chemical modification of arginine residues and mass spectrometric analysis. *Biochemistry* 41(6):1759–1766
104. Hochleitner EO, Borchers C, Parker C, Bienstock RJ, Tomer KB (2000) Characterization of a discontinuous epitope of the human immunodeficiency virus (HIV) core protein p24 by epitope excision and differential chemical modification followed by mass spectrometric peptide mapping analysis. *Protein Sci* 9(3):487–496
105. Xu GH, Chance MR (2007) Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. *Chem Rev* 107(8):3514–3543
106. Konermann L, Stocks BB, Pan Y, Tong X (2010) Mass spectrometry combined with oxidative labeling for exploring protein structure and folding. *Mass Spectrom Rev* 29(4):651–667
107. Roeser J, Bischoff R, Bruins AP, Permentier HP (2010) Oxidative protein labeling in mass-spectrometry-based proteomics. *Anal Bioanal Chem* 397(8):3441–3455

108. Kaur P, Kiselar JG, Chance MR (2009) Integrated algorithms for high-throughput examination of covalently labeled biomolecules by structural mass spectrometry. *Anal Chem* 81(19):8141–8149
109. Mendoza VL, Vachet RW (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev* 28(5):785–815
110. Hambly D, Gross M (2005) Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *J Am Soc Mass Spectrom* 16(12):2057–2063
111. Gau BC, Sharp JS, Rempel DL, Gross ML (2009) Fast photochemical oxidation of protein footprints faster than protein unfolding. *Anal Chem* 81(16):6563–6571
112. Suhnel J (2001) Beyond nucleic acid base pairs: from triads to heptads. *Biopolymers* 61(1):32–51
113. Balasubramanian S, Hurley LH, Neidle S (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov* 10(4):261–275
114. Krishnan Y, Simmel FC (2011) Nucleic acid based molecular devices. *Angew Chem Int Ed* 50(14):3124–3156
115. Lightwahl KJ, Springer DL, Winger BE, Edmonds CG, Camp DG, Thrall BD, Smith RD (1993) Observation of a small oligonucleotide duplex by electrospray ionization mass spectrometry. *J Am Chem Soc* 115(2):803–804
116. Goodlett DR, Camp DG, Hardin CC, Corregan M, Smith RD (1993) Direct observation of a DNA quadruplex by electrospray ionization mass spectrometry. *Biol Mass Spectrom* 22(3):181–183
117. Rosu F, Gabelica V, Poncelet H, De Pauw E (2010) Tetramolecular G-quadruplex formation pathways studied by electrospray mass spectrometry. *Nucleic Acids Res* 38(15):5217–5225
118. Rosu F, De Pauw E, Gabelica V (2008) Electrospray mass spectrometry to study drug-nucleic acids interactions. *Biochimie* 90(7):1074–1087
119. Brodbelt JS (2010) Evaluation of DNA/ligand interactions by electrospray ionization mass spectrometry. *Annu Rev Anal Chem* 3(1):67–87
120. Millard JT (1999) Molecular probes of DNA structure. *Compr Nat Prod Chem* 7:81–103
121. Smith SI, Brodbelt JS (2010) Rapid characterization of cross-links, mono-adducts, and non-covalent binding of psoralens to deoxyoligonucleotides by LC-UV/ESI-MS and IRMPD mass spectrometry. *Analyst* 135(5):943–952
122. Parr C, Pierce SE, Smith SI, Brodbelt JS (2011) Investigation of the reactivity of oligodeoxynucleotides with glyoxal and KMnO₄ chemical probes by electrospray ionization mass spectrometry. *Int J Mass Spectrom* 304(2–3):115–123
123. Gmeiner WH, Sahasrabudhe P, Pon RT (1995) Use of shaped pulses for semi-selective excitation of imino H-1 resonances in duplex DNA and RNA. *Magn Reson Chem* 33(6):449–452
124. Nonin S, Jiang F, Patel DJ (1997) Imino proton exchange and base-pair kinetics in the AMP-RNA aptamer complex. *J Mol Biol* 268(2):359–374
125. Li TS, Johnson JE, Thomas GJ (1993) Raman dynamic probe of hydrogen-exchange in bean pod mottle virus—base-specific retardation of exchange in packaged ssRNA. *Biophys J* 65(5):1963–1972
126. Hofstadler SA, Sannes-Lowery KA, Griffey RH (2000) Enhanced gas-phase hydrogen–deuterium exchange of oligonucleotide and protein ions stored in an external multipole ion reservoir. *J Mass Spectrom* 35(1):62–70
127. Hofstadler SA, Sannes-Lowery KA, Hannis JC (2005) Analysis of nucleic acids by FTICR MS. *Mass Spectrom Rev* 24(2):265–285
128. Yu E, Fabris D (2003) Direct probing of RNA structures and RNA-protein interactions in the HIV-1 packaging signal by chemical modification and electrospray ionization Fourier transform mass spectrometry. *J Mol Biol* 330(2):211–223
129. Kellersberger KA, Yu E, Kruppa GH, Young MM, Fabris D (2004) Top-down characterization of nucleic acids modified by structural probes using high-resolution tandem mass spectrometry and automated data interpretation. *Anal Chem* 76(9):2438–2445

130. Turner KB, Yi-Brunozzi HY, Brinson RG, Marino JP, Fabris D, Le Grice SFJ (2009) SHAMS: combining chemical modification of RNA with mass spectrometry to examine polypurine tract-containing RNA/DNA hybrids. *RNA* 15(8):1605–1613
131. Fabris D, Yu ET (2010) Elucidating the higher-order structure of biopolymers by structural probing and mass spectrometry: MS3D. *J Mass Spectrom* 45(8):841–860
132. Kaltashov IA, Eyles SJ (2012) *Mass spectrometry in structural biology and biophysics: architecture, dynamics, and interaction of biomolecules*, 2nd edn. Wiley, Hoboken, NJ
133. Bobst CE, Kaltashov IA (2011) *Advanced mass spectrometry-based methods for the analysis of conformational integrity of biopharmaceutical products*. *Curr Pharm Biotechnol* 12(10):1517–1529
134. Fabris D (2010) A role for the MS analysis of nucleic acids in the post-genomics age. *J Am Soc Mass Spectrom* 21(1):1–13